
テレビ向け
シーン検索技術及び番組推薦技術

廣 井 和 重

電気通信大学大学院情報システム学研究科

博士(工学)の学位申請論文

2015年 6 月

テレビ向け
シーン検索技術及び番組推薦技術

電気通信大学大学院情報システム学研究科

博士(工学)の学位申請論文

博士論文審査委員会

主査	森田	啓義	教授
委員	加藤	聰彦	教授
委員	吉永	努	教授
委員	田野	俊一	教授
委員	栗原	聡	教授

著作権所有者

廣 井 和 重

2015年

Scene Search Technology and Program Recommendation Technology for Television Systems

Kazushige Hiroi

Abstract

In recent years, we can enjoy a lot of television programs via digital terrestrial television broadcasting, BS (Broadcasting Satellite) digital television broadcasting, CS (communication satellite) digital television broadcasting, and so on. And the capacity for the storage such as a hard disk etc. is also turning into large scale. Moreover, Motion-picture-compression technologies, such as MPEG (Moving Picture Experts Group) and H.26x, have also evolved. With these backgrounds, the amount of television programs which a user holds in their television systems and recorder systems is becoming huge. However, even if the amount of enjoyable television programs increase, the time for users is limited in 24 hours a day and they do not have enough time to enjoy all of them. Rather, time to watch and enjoy these television programs is decreasing because users have many things to do and there is a lot of information to process because of much

information recently. So, even if a lot of television programs can be recorded, there might be no time to watch all of them. And as the choices to television programs become a lot, it becomes difficult to select the television programs to be watched actually. Therefore, the case that users can not watch their favorite television programs increases because they can not select them after all. This will lead the case increase that users will forget to record or watch their favorite television programs, even if they have recorder systems. So there are a lot of needs to watch and enjoy a lot of favorite television programs efficiently.

In this thesis, we propose the “Scene search function” which enables users to enjoy only favorite scenes in television programs in order to enable users to watch more interested television programs from a lot of television programs in limited time. Also, we propose the “Television program recommender function” which enables recorder systems to recommend user’s favorite television programs or to record them automatically in order to enable users to watch and select their favorite programs easily without missing and forgetting them. Especially, we propose technologies for adding these functions to television systems which have recorder function.

The television system must be the accessible product for all users. Therefore, requirements to the cost are very stringent. So there are a lot of limitations for processing power and user-interface in the television system. And requirements to the usability and the accuracy of function which has an influence on the reliability are also stringent within these limitations.

Especially, in this research, our target television system has the cheap CPU which operates 300MHz and the cost effective memory whose capacity is 256MByte. So, for the “Scene search function”, our proposed technology must consume fewer than 3MHz of CPU load and fewer than 0.5MByte of memory for adding this function to television systems which have a cheap CPU and a cost effective memory. And it must resolve the mismatch problem between user specified scenes and start positions of topics for specified scenes which might be a problem at the time of a scene search. And for the “Television program recommender function”, our proposed technology must finish recommendation process for television programs within 10 seconds with these limited hardware resources because of short response time. And it must optimize balance of recall and precision for recommendation, which has not been considered so far, in order to recommend user’s favorite television programs in proper quantities and qualities as far as it's possible.

In order to resolve these problems, for the “Scene search function”, we propose methods for scene indexing with the scene keyword extraction method by practical use of subtitle information included in television programs. Especially, we propose the multi-indexing method for scenes with two types of dictionaries. One is the fixed phrase dictionary which contains common keywords for each genre of a television program. Another is the peculiar phrase dictionary which contains peculiar keywords for each television program. The mismatch problem between users specified scenes and start positions of topics for

specified scenes which might be a problem at the time of a scene search can be resolved with these two types and different level of dictionaries. We also propose an efficient scene expression by coding keywords in the scene with these dictionaries. The conventional technology consumes 8MHz of CPU load and 2MByte of memory for scene indexing. And F-measure for scene search is 65.9% (Recall is 80% but Precision is 56%). On the other hand, proposed method consumes only 1.835MHz of CPU load and 0.384MByte of memory for scene indexing under the same condition. And F-measure for scene search can be 81.5% (Recall can be 83% and Precision can be 80%). So, the proposed method enables us to add the “Scene search function” to the television system which has the cheap CPU and the cost effective memory.

For the “Television program recommendation function”, we define “Playback”, “Reservation”, “Search”, and “Record” operation as user’s operations to the television system which are minimum necessary and express user’s preferences exactly. And also, we define “Program title”, “Program keywords”, “Program genre”, “Broadcasted channel”, and “Broadcasting time” as important attributes for television programs which are also minimum necessary and express user’s preferences exactly. And we define weights for these parameters. The proposed method can calculate recommendation scores for television programs based on “contents based recommendation methods” with these weights for parameters. Furthermore, we define the “Number-Score curve (N-S Curve)” which indicates the relationship between the recommendation score and the number of recommended items. So,

proposed method can optimize recall and precision for the television program recommendation by controlling the number of recommended television programs using the N-S Curve. The conventional technology takes 26 seconds for recommendation process for 3000 television programs, which are television programs for one week in Japanese digital terrestrial television broadcasting, with the recordable television system which have a 300MHz CPU and a 256MByte memory. And recall and precision for television program recommendation are 62% and 71% respectively (F-measure is 66.2%). On the other hand, the proposed method takes only 8 seconds for recommendation process under the same condition. And recall and precision for television program recommendation can be 78% and 82% respectively (F-measure can be 79.9%). So, the proposed method can improve balance of recall and precision for the television program recommendation and enables us to add the “Television program recommendation function” to the television system which has the cheap CPU and the cost effective memory.

Furthermore, we inquire the possibility of the recommendation with social-media data and suggest the recommendation method considering world trends predicted by analyzing social-media data.

Proposed technologies enable users to watch and enjoy efficiently not only television programs but also scenes in their programs with television systems which have restrictions for the processing power and a user-interface. So people who are becoming busy in recent years can enjoy a lot of television

programs in short time with low cost television systems by
proposed functions and technologies.

テレビ向け シーン検索技術及び番組推薦技術

廣 井 和 重

論文概要

近年，地上波デジタル放送，BS(Broadcasting Satellite)デジタル放送，CS(Communication Satellite)デジタル放送等により，視聴可能なテレビ番組が増えている．加えて，HDD(Hard Disk Drive)の大容量化及びMPEG(Moving Picture Experts Group)やH.26x等での動画圧縮技術も進化している．こうした背景から，ユーザが保有するテレビやレコーダでのテレビ番組の量は膨大になっている．一方，視聴可能なテレビ番組がどんなに多くなっても，1日24時間という限られた時間の中でユーザがこれらのテレビ番組を視聴できる時間には限りがある．むしろ，行動の多様化や情報の大量化に伴い，ユーザがゆっくりとテレビ番組を楽しめる時間は減り，折角大量のテレビ番組を録画しても，これらをすべて観る時間はなくなっている．また，視聴したいテレビ番組に対する選択肢が多くなった分，実際に視聴する際のテレビ番組の選択に時間がかかり，結局は観たい番組を選択できず，これらを観ることができないケースも増えている．さらに，見忘れ等も多くなり，折角録画機器があっても録画を忘れるケースも増えている．

本研究では、これらの課題を解決し、大量のテレビ番組の中からユーザが興味ある番組を短時間で視聴可能とするために、ユーザが好むシーンのみを視聴可能とする「シーン検索機能」を実現する。また、テレビ番組をユーザが簡単かつ見落としなく視聴可能とするために、ユーザの好みに合ったテレビ番組を推薦及び自動録画可能とする「テレビ番組推薦機能」を実現する。特に、これらの機能を録画機能付きのテレビに付加可能とする。

テレビはユーザにとって身近な製品とするため、コストに対する要求が厳しく、処理能力とユーザインターフェースに制限がある。また、このような制限がある中で、信頼性に影響を与える機能の精度や使い勝手に対する要求も厳しい。本研究では、特に、安価な300MHzのCPU、256MByteのメモリを搭載した普及型の録画機能付きテレビにおいてこれらの機能を付加可能とする。

シーン検索機能では、テレビ番組の録画時に余っている計算リソースでシーンのインデクシングを可能とするため、CPU負荷3MHz、使用メモリ量0.5MByte以下でシーンのインデクシング可能な処理負荷とする。また、シーン検索時に問題となる「指定シーンとトピック開始位置のズレ問題」を解消する。テレビ番組推薦機能では、ユーザがテレビに対して何らかの処理要求を行った場合、結果が出てくるまでに待つことができる時間は長くて10秒程度であるため、本テレビ番組推薦機能では10秒以下で推薦処理可能とする。また、ユーザの好む番組を可能な限り過不足なく推薦するため、推薦適合率と再現率のバランスを最適化する。

これを実現するために、シーン検索機能向けには、番組に付随する字幕情報を利用することでシーンキーワードを抽出し、シーンのインデクシングを行う方式を提案する。特に、テレビ番組のジャンル毎に共通のキーワードからなる定型句辞書と、番組に固有のキーワードからなる動画固有辞書による多重インデクシング方式を提案する。これら二種類の意味レベルの異なる辞書により、指定シーンとトピック開始位置のズレ問題を解決する。また、これらの辞書を用いたキーワードの符号化によるシーンの効率的な表現を提案

する．従来技術では，インデクシング処理にCPU負荷として8MHz，使用メモリ量として2MByteを要し，シーン検索精度を示すF値は65.9%(適合率56%，再現率80%)である．これに対して提案方式では，インデクシングに必要となるCPU負荷を1.835MHz，必要メモリ量を0.384MByteとすることができる．また，シーン検索におけるF値を81.5%(適合率80%，再現率83%)とすることができる．これにより，普及型の録画機能付きテレビにおいて，CPUやメモリを追加あるいは変更することなくシーン検索機能を付加可能とした．

テレビ番組推薦機能向けには，ユーザの嗜好を的確に表現する必要最低限のテレビ操作として，再生，予約，検索，録画を定義する．また，テレビ番組の重視属性として，番組タイトル，番組キーワード，番組ジャンル，放送チャンネル，放送時間を定義し，これらの重み付けにより推薦対象のテレビ番組に対するユーザの嗜好度を示す推薦スコアの計算方式を提案する．さらに，推薦スコアと推薦番組数の関係を示す“Number-Score Curve (N-S Curve)”を考案し，推薦番組数を制御することで，テレビ番組推薦における適合率と再現率のバランスを最適化する．従来技術では，300MHzのCPU，256MByteのメモリを搭載した普及型の録画機能付きテレビにおいて，一週間の番組(3000番組)に対する推薦処理に26秒を要し，推薦適合率と再現率は，それぞれ71%及び62%(F値=66.2%)である．これに対して提案方式では，同様の条件下で推薦処理を8秒で実施可能とし，テレビ番組推薦における適合率と再現率をそれぞれ82%及び78%(F値=79.9%)とすることができる．これによりテレビ番組推薦機能における適合率と再現率のバランスを改善し，普及型の録画機能付きテレビにおいて，CPUやメモリを追加あるいは変更することなくテレビ番組推薦機能を付加可能とした．

さらに，本論文では，近年増加しているソーシャルメディアデータを活用した推薦の可能性についても検討し，世の中のトレンドを反映した推薦が可能であることを示す．

これらにより，処理能力及びユーザインターフェースに制限があるテレビ

において、好きなテレビ番組を見落としなく素早く選択して、そのシーンを短時間で視聴可能とし、近年多忙化する生活者が低価格なテレビを用いて、短時間で大量のテレビ番組を楽しむことを可能とした。

目次

第1章 序論.....	1
1.1 本研究の背景と目的	2
1.2 本研究の必要性	3
1.3 本研究の課題と目標	6
1.4 本論文の構成	9
第2章 テレビにおける付加機能実現とは.....	12
2.1 はじめに	13
2.2 テレビの概要	14
2.3 設計における要求	15
2.4 技術的な主眼点と問題意識	17
2.4.1 処理負荷.....	17
2.4.2 精度.....	21
2.5 処理負荷と精度に関する位置づけ	22
2.6 短時間視聴に関する従来の取り組み	23
2.7 第2章まとめ	26
第3章 シーン検索技術.....	28
3.1 はじめに	29
3.2 関連技術	30
3.3 提案方式	37
3.3.1 字幕情報.....	38
3.3.2 定型句辞書.....	39
3.3.3 動画固有辞書	40
3.3.4 多重インデクシング法.....	42
3.4 テレビ向けシーン検索機能の実現	43
3.4.1 字幕符号化方式.....	44
3.4.2 字幕処理ルール.....	47
3.4.3 辞書構築方法	56

3.4.4 シーン検索機能実現エンジン.....	57
3.5 提案技術の評価と関連技術との比較.....	59
3.5.1 提案技術の評価.....	60
3.5.2 関連技術との比較.....	62
3.5.3 評価結果及び比較結果に対する考察.....	65
3.6 新規性と有用性	68
3.7 第3章のまとめ	70
第4章 テレビ番組推薦技術.....	71
4.1 はじめに	72
4.2 関連技術	73
4.3 提案方式	83
4.3.1 Number-Score Curve (N-S Curve)	84
4.3.2 N-S Curve における推薦精度の関係	86
4.3.3 推薦アイテム数決定方法	88
4.4 テレビ番組推薦への適用	91
4.5 提案技術の評価と関連技術との比較.....	103
4.5.1 提案技術の評価.....	103
4.5.2 関連技術との比較.....	105
4.5.3 評価結果及び比較結果に対する考察.....	109
4.6 新規性と有効性	110
4.7 第4章まとめ	112
第5章 今後の展望.....	114
5.1 はじめに	114
5.2 ソーシャルメディア環境の現状.....	116
5.2.1 ソーシャルメディアの利用状況	116
5.2.2 ソーシャルメディアの利用目的	117
5.2.3 ユーザの利用機器と行動の変化	117
5.2.4 市場への影響	118
5.2.5 ソーシャルメディア活用のメリットとデメリット	118
5.3 ソーシャルメディアデータ利活用の可能性.....	119
5.4 トренд把握方式の検討.....	121
5.4.1 検討対象データ.....	121
5.4.2 流行る可能性のあるアイテムの特定.....	122
5.4.3 トренд変動要因の把握.....	125
5.4.4 トрендを反映した需要予測.....	126

5.5 トренд予測結果の評価	127
5.6 第5章のまとめ	128
第6章 まとめ.....	131
6.1 結言.....	131
6.1.1 シーン検索技術.....	134
6.1.2 テレビ番組推薦技術.....	134
6.2 今後の展開.....	136
謝辞.....	139
関連論文の印刷公表の方法及び時期	141
参考論文の印刷公表の方法及び時期	142
その他の発表及び講演	143
関連主筆出願特許	144
参考文献.....	147
著者略歴	158

図目次

- 図1-1 録画番組に対する視聴実態
- 図1-2 シーン検索機能に対するニーズ
- 図1-3 テレビ番組の番組推薦機能及び自動録画機能に対するニーズ
- 図1-4 シーン検索及びテレビ番組推薦/自動録画機能に対する意見
- 図2-1 録画時の計算リソース使用率に関する設計思想
- 図2-2 録画時のCPU利用率
- 図2-3 処理負荷と精度の位置づけ
- 図3-1 シーン検索機能の実現目標
- 図3-2 シーン検索機能を実現する技術
- 図3-3 提案方式における処理の流れ
- 図3-4 **Transport Stream**
- 図3-5 多重インデクシング方式の構成
- 図3-6 多重インデクシングの例
- 図3-7 インデクシングと時刻の補正
- 図3-8 SI情報とキーワード
- 図3-9 シーン検索機能実現エンジンの構成
- 図3-10 シーン検索機能のユーザインターフェース
- 図3-11 関連方式との比較結果
- 図3-12 処理負荷と精度の位置づけに対するマッピング
- 図3-13 野球のキーワードに対するシーン検出精度改善過程
- 図4-1 テレビ番組推薦機能の実現目標
- 図4-2 テレビ番組推薦機能を実現する技術

- 図4-3 提案方式確立に向けたアプローチ
- 図4-4 推薦スコア
- 図4-5 N-S Curveの例
- 図4-6 推薦スコアでソートしたアイテム
- 図4-7 N-S Curveにおける変曲点
- 図4-8 適合率と再現率
- 図4-9 推薦アイテム数と適合率及び再現率の関係
- 図4-10 推薦アイテム数決定推薦閾値の選定
- 図4-11 推薦アイテム数の決定
- 図4-12 ユーザごとの推薦アイテム数の決定
- 図4-13 ユーザ嗜好データベース
- 図4-14 推薦番組の決定
- 図4-15 推薦計算例におけるN-S Curve
- 図4-16 推薦度の決定方法
- 図4-17 推薦計算例における推薦度の決定
- 図4-18 推薦結果および推薦度の表示例
- 図4-19 関連方式との比較結果
- 図4-20 処理負荷と精度の位置づけに対するマッピング
- 図4-21 推薦閾値の変化に対する推薦適合率，再現率，及びF値の変化
- 図5-1 定番アイテムにおける売上額とTwitter発言数の時系列推移
- 図5-2 新商品における売上額とTwitter発言数の時系列推移
- 図5-3 商品発売前のTwitter発言数と売上の関係
- 図5-4 新商品についての売上とTwitter上での評判の関係
- 図5-5 新商品の需要予測結果例
- 図5-6 売上推移の予測結果
- 図5-7 需要予測誤差の評価
- 図5-8 ソーシャルグラフ及びインタレストグラフによる推薦

図6-1 処理負荷と精度の位置づけにおける今後のターゲット

図6-2 未来志向型の推薦

表目次

表2-1	推薦アイテム数に関するトレードオフ
表3-1	シーン検索関連技術の本研究目標達成に対する近さ
表3-2	定型句辞書の例
表3-3	動画固有辞書の例
表3-4	字幕特徴データベース
表3-5	動画固有辞書によるインデクシング
表3-6	定型句辞書によるインデクシング
表3-7	インデックスデータ
表3-8	チャプタ位置及び提示キーワード
表4-1	推薦方式に関する関連技術のまとめ
表4-2	適合率と再現率のバランス及び処理負荷に関する関連技術のまとめ
表4-3	ジャンル適応キーワード
表4-4	嗜好抽出に用いる属性と重み
表4-5	嗜好抽出に用いるテレビ操作と重み
表4-6	属性に対する重みの例
表4-7	操作に対する重みの例
表4-8	キーワードテーブルの例
表4-9	ジャンルテーブルの例
表4-10	テレビ操作ログ
表4-11	テレビ操作ログの例
表4-12	ユーザ嗜好データベースにおけるタイトルテーブル
表4-13	ユーザ嗜好データベースにおけるジャンルテーブル

表 4-14	ユーザ嗜好データベースにおけるキーワードテーブル
表 4-15	ユーザ嗜好データベースにおけるチャンネルテーブル
表 4-16	ユーザ嗜好データベースにおける放送時間テーブル
表 4-17	放送時間 (Time) の表現
表 4-18	推薦対象の番組と推薦スコアの計算結果
表 4-19	推薦表示すべきテレビ番組の決定結果
表 4-20	推薦計算例による推薦番組および推薦度の結果

第1章

序論

テレビは市場に登場して以来、多くの機能が取り入れられてきた。茶の間のテレビは白黒からカラーへと変わり、やがてリモコンにより操作できるようになった。また、テレビの低価格化に伴い、それまで茶の間に置かれていたテレビは、家庭の各部屋に置かれるようになった。それだけでなく、ポータブル化の波に乗り、テレビもポータブル化され、屋外にテレビを持ち出すことも可能となった。周辺装置としても、VTR(Video Tape Recorder)やHDR(Hard Disk Recorder)等が登場し、一般家庭でテレビ番組を容易に録画可能となった。これらの背景に伴い、消費者のテレビの視聴スタイルも徐々に変わっていった。テレビが茶の間に1台しかなかった時代では、皆が茶の間に集まり、テレビ付近でチャンネルを操作しながらテレビ番組を視聴していた。リモコンの登場により、カウチでポテトを食べながら番組を視聴するスタイルへと変わっていった。家庭の各部屋にテレビが置かれるようになると、個人が別個にテレビを占有し、番組を楽しむようになった。また、ポータブルテレビにより電車や車の中など、外出先でも番組を楽しめるようになった。VTRやHDR等が登場すると、テレビ番組を録画して、時間に縛られず好きな時間にテレビ番組を楽しむようになった。現在、テレビはさらに進化を遂げ、デジタル放送により、高画質化及び多チャンネル化された番組を楽しむことができる。また、ネットワークインフラの発展と整備により、ネットワー

クを使用してテレビ番組を視聴できるようになり、いわゆるビデオオンデマンドを始めとするネット動画の視聴も可能となっている。今後は3D(dimension)4K/8K解像度といった臨場感のある番組も視聴できるようになり、これらの大量かつ魅力的なテレビ番組をこれまでのように放送局主導のプッシュ型ではなく、ユーザ主導でプル型かつインタラクティブに楽しむことができるようになると考えられる。この環境では、テレビの新視聴スタイルとして、大量かつ魅力的なテレビ番組を見落とし無く、短時間で視聴するニーズが増えるものと考えられる。本研究ではこれらのニーズを充足するための機能と、その実現技術を提案する。特に、ユーザが好みのシーンを検索して視聴可能とする「シーン検索機能」と、ユーザの好みに合った番組を推薦/自動録画可能とする「テレビ番組推薦機能」を提案し、これらの機能を低価格で普及型の録画機能付きテレビに付加可能とするための技術を提案する。

1.1 本研究の背景と目的

近年、地上波デジタル放送、BS (Broadcasting Satellite) デジタル放送、CS (Communication Satellite) デジタル放送等により、視聴可能なテレビ番組が増えている。加えて、ハードディスクやブルーレイディスク等の記憶装置も大容量化がなされ、MPEG (Moving Picture Experts Group) やH.26x等での動画圧縮技術も進化している。こうした背景から、ユーザは大量のテレビ番組をHDR (Hard Disk Recoder) やBDR (Blu-ray Disc) 等に録画できるようになり、今後はこれまでのように放送局主導のプッシュ型ではなく、ユーザ主導で大量のテレビ番組をプル型かつインタラクティブに楽しむことができるようになると考えられる。このような環境ではユーザが「観たいテレビ番組を大量かつ気軽に録画し、好きな時間に視聴する」テレビ視聴スタイルになると考えられ、ユーザにとっては益々視聴したいテレビ番組が増大になると考えられる。

しかしながら、視聴可能なテレビ番組がどんなに多くなっても、1日24時間という限られた時間の中で、ユーザがテレビ番組を視聴可能な時間には限りがある。むしろ、行動の多様化や情報の大量化に伴い、ゆっくりとテレビ番組を楽しむ時間は減り、折角大量のテレビ番組を録画しても、これらをすべて観る時間はなくなってくると考えられる。また、視聴したいテレビ番組に対する選択肢が多くなった分、実際に視聴する際のテレビ番組の選択に時間がかかり、結局は観たい番組を選択できず、観ることができないケースも増える。好みの番組を見つけることができないケースや見忘れも多くなり、折角録画機器があっても録画を忘れることも多くなると考えられる。

そこで、本研究では、これらの課題を解決し、大量のテレビ番組の中からユーザが興味ある番組を見落とし無く、短時間で視聴可能とする機能を実現する。特に、ユーザが好むシーンのみを視聴可能とする「シーン検索機能」を実現する。また、シーンだけでなくテレビ番組そのものをユーザが簡単かつ見落としなく視聴可能とするために、ユーザの好みに合ったテレビ番組を推薦/自動録画する「テレビ番組推薦機能」を実現する。単に実現するだけでなく、これらの機能を普及型のテレビに付加可能とする。普及型のテレビはコストに対する要求が厳しく、処理能力とユーザインターフェースに制限がある。また、このような制限がある中で、信頼性に影響を与える機能の精度や使い勝手に対する要求が厳しい。本研究成果により、消費者にとって身近なテレビにおいて、好きなテレビ番組を見落としなく素早く選択して、好きなシーンのみを選択して視聴可能とする。そして、近年多忙化する生活者が低価格なテレビを用いて、短時間で大量のテレビ番組を楽しむことを可能とする。

1.2 本研究の必要性

前述のとおり、視聴可能なテレビ番組が多くなっても、これらをゆっくり

と楽しむ時間は減ってきている。図1-1は、著者らが録画機器ユーザに対してヒヤリングした結果であり、ユーザによるテレビ番組の録画実態及び録画番組に対する視聴実態を示している。

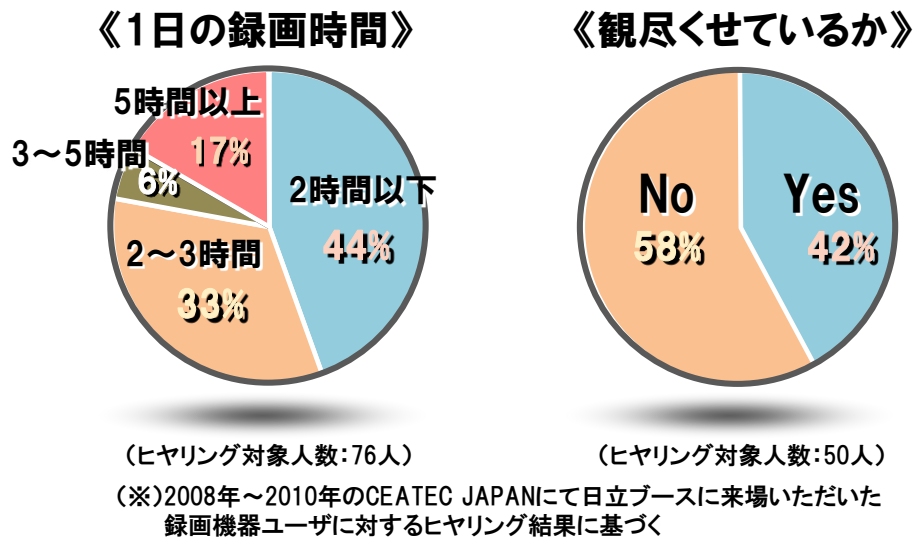


図1-1 録画番組に対する視聴実態

本ヒヤリング結果では、1日のテレビ録画時間が2時間以下のユーザで44%、2時間以上録画しているユーザが56%であった。一方、これらの録画番組を見尽くせているユーザは42%であり、58%のユーザがこれらの録画番組を観尽くせていない。この影響もあり、録画番組を短時間で視聴するニーズが高まってきている。

図1-2は、テレビ番組のシーン検索機能に対するニーズのヒヤリング結果である。本ヒヤリング結果では、シーン検索機能を「欲しい」と答えたユーザは68%であり、「あれば嬉しい」と答えたユーザを含めると、そのニーズは84%となっている。

また、視聴したいテレビ番組に対する選択肢が多くなった分、実際に視聴する際のテレビ番組の選択に時間がかかり、結局は観たい番組を選択できず、観ることができないケースも増えている。見落としや見忘れも多くなり、折

角録画機器があっても録画することを忘れるケースも多くなっている。そのため、テレビ番組の番組推薦機能及び自動録画機能に対するニーズが高まってきている。

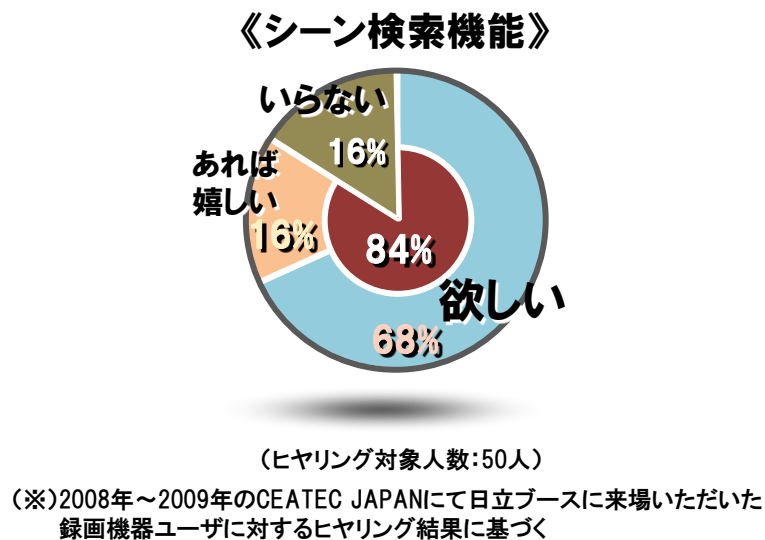


図1-2 シーン検索機能に対するニーズ

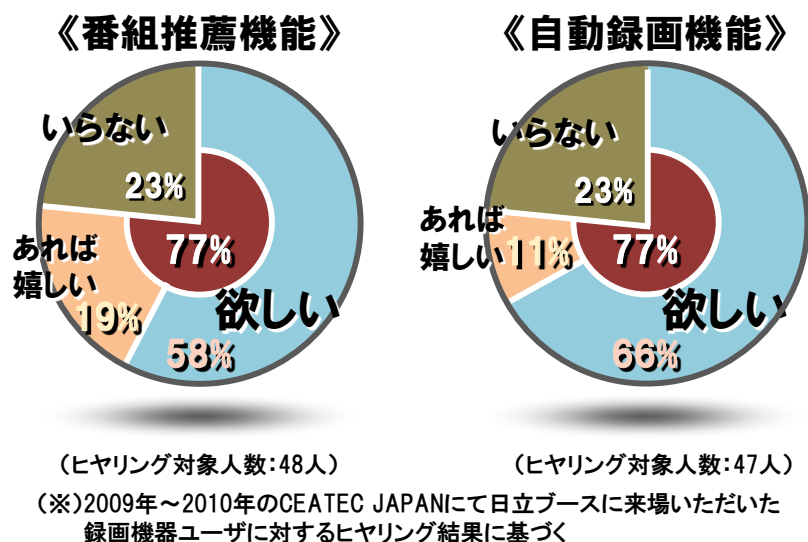


図1-3 テレビ番組の番組推薦機能及び自動録画機能に対するニーズ

図1-3は、テレビ番組の番組推薦機能及び自動録画機能に対するニーズのヒヤリング結果である。本ヒヤリング結果では、これらの機能を「欲しい」と答えたユーザは約6割、「あれば嬉しい」と答えたユーザを含めると、そのニーズは約8割に達している。特に、自動録画機能は、ユーザにとって分かりやすく、録画し忘れ防止や録画操作の手間が省ける等、メリットが明確であるため、多くのユーザが明確に欲しいと回答した。

これらの結果から、大量のテレビ番組の中からユーザが興味ある番組を見落とし無く、短時間で視聴可能とする機能の必要性は高いといえる。

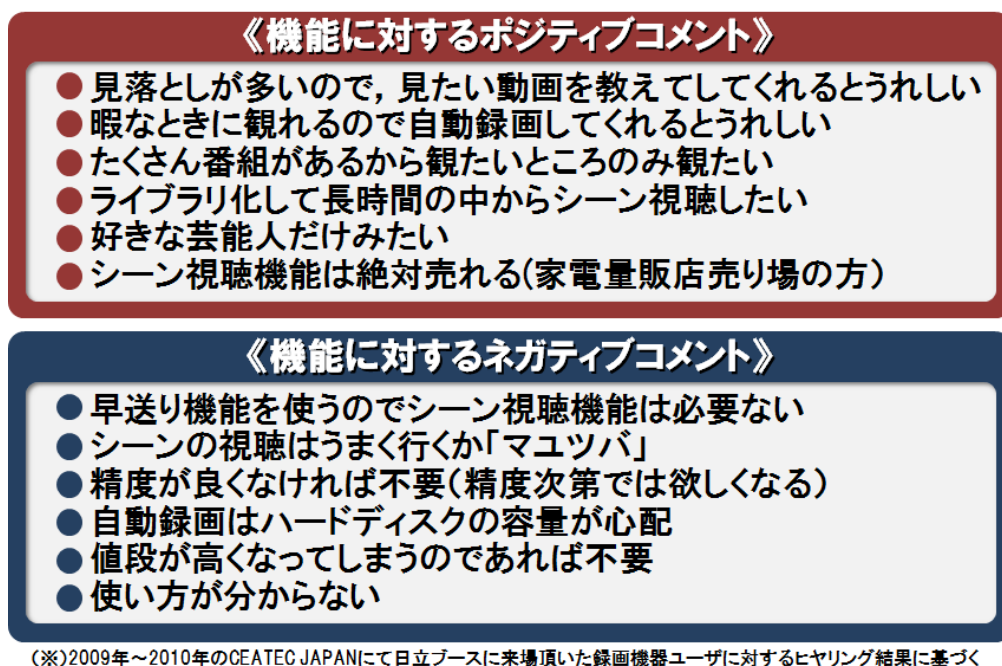


図1-4 シーン検索及びテレビ番組推薦/自動録画機能に対する意見

1.3 本研究の課題と目標

前節で述べたシーン検索機能やテレビ番組推薦機能及び自動録画機能について消費者からフリーな意見をヒヤリングしたところ、これらの機能に対し

ではポジティブな意見とネガティブな意見があった。図1-4は、著者らが録画機器ユーザに対してヒヤリングした結果である。本ヒヤリング結果からわかるとおり、ネガティブな意見の多くは、コストや精度及び使い勝手に関するものが多く、これらを改善することが重要となる。

一方、テレビは一般家庭に普及して以来、子供からお年寄りまで、誰もが簡単に楽しめる家庭の中心的存在である。その身近さから、テレビの設計におけるコストや信頼性及び使い勝手に対する要求仕様は、PC (Personal Computer) 等の他のデジタルメディア機器に比べて非常に厳しい。ユーザインターフェースはリモコンが主体であり、ボタンによる操作が前提となっている。リモコンで離れた場所からテレビを操作することも多く、視覚上の表示解像度が低い。CPUやメモリも可能な限り低価格のものを採用しており、余剰能力を持たない。このため、テレビはコストを最適化した設計になっており、最大処理負荷時の処理能力に余裕がなく、他の処理をほとんど行えない設計となっている。一方で、今後も新しいテレビを市場に投入し、より良いコンセプトを消費者に提供するために付加機能を実現する必要がある。しかし、テレビに新たな付加機能を実現する場合には、価格を抑えるために設計変更を少なくし、CPUやメモリを変えないことも多い。また、ユーザに長く安心して使用してもらえる機能とするために、使い勝手や精度の良い信頼性のある機能とすると共にリモコンで簡単に扱える機能とする必要がある。特に、シーン検索機能では、通常、キーワードや顔画像が指定された場合に、そのキーワードや顔が出現するシーンを検出して視聴可能とする。このため、指定されたシーンに関するトピックの開始位置からズレた位置から再生されることが多く、ユーザにとっては精度が良い検索とは言えない。テレビ番組推薦機能では、推薦番組数を多くすると、ユーザは多くの選択肢を得ることが可能となるが、好みではない番組も推薦される可能性があり、所望の番組を選択することが困難となる。一方で、推薦番組数が少ない場合には、ユーザは好みの番組を選択することが容易となるが、選択肢が少なく、所望の番組が

推薦されない可能性がある。すなわち、これらはトレードオフであり、視覚上の解像度が低い画面に表示された推薦番組をリモコンで簡単に選択可能とするには、最適な推薦番組数でこれらのバランスをとることが重要となる。

そのため、本研究では、以下の主眼点と問題意識に基づいてシーン検索機能とテレビ番組推薦機能をテレビに付加可能とする。

<主眼点>

- (1) シーン検索機能とテレビ番組推薦機能を低コストに実現する。具体的には、普及型の録画機能付きテレビにおいて、CPUやメモリを変更或いは追加することなく、これらの機能を付加可能とする。
- (2) 高精度なシーン検索機能とテレビ番組推薦機能を実現する。具体的には、シーン検索においては、ユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とする。また、テレビ番組の推薦において最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とする。

<問題意識>

上記の主眼点で述べた項目を実現しないとコストや信頼性及び使い勝手に対する要求仕様が厳しく、処理能力やユーザインターフェースに制限があるテレビにおいて、目的とするシーン検索機能とテレビ番組推薦機能を付加することが困難となる。特に、主眼点(1)の項目については、普及型のテレビではCPU及びメモリは全体の1%ほどしか使えず、これ以下でシーン検索機能を実現できるようにする必要がある。また、テレビ番組推薦機能では、普及型のテレビに備わっている少ないCPU及びメモリ資源で推薦処理を高速に実施し、リクエストに対してユーザが待つことができると考えられる10秒以下で推薦結果を提示できるようにする必要がある。これらを実現しなければ、CPUやメモリを高速なものに変更或いは追加する必要があり、テレビの原価コストが

上がってしまう。このため、普及型のテレビにシーン検索機能及びテレビ番組推薦機能を付加することが許容されない可能性がある。また、主眼点(2)の項目については、信頼性がなく使い勝手の良くない機能であると判断され、ユーザに使用されない機能となる可能性があり、シーン検索機能及びテレビ番組推薦機能をテレビに付加するという判断にはなりえない可能性がある。

1.4 本論文の構成

本論文の構成は以下の通りである。

第2章では、テレビにおける付加機能実現に対する要件について述べる。特に、テレビにシーン検索機能及びテレビ番組推薦機能を付加する場合の設計要求について、その背景を踏まえて説明する。また、テレビにシーン検索機能とテレビ番組推薦機能を付加する場合の技術的な主眼点と問題意識を説明した上で、実現すべきシーン検索機能とテレビ番組推薦機能の要求仕様を明確化する。具体的には、300MHzのCPU、256MByteのメモリを持つ普及型の録画機能付きテレビにおいて、シーン検索機能をCPU負荷3MHz、使用メモリ量0.5MByte以下で実現する必要があること、及び、テレビ番組の推薦処理を10秒以下で完了する必要があることを示す。また、シーン検索においては、ユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とし、テレビ番組の推薦において最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とする必要があることを示す。

第3章では、第2章で述べた要件を満たすシーン検索機能の実現技術として、番組に付随する字幕情報を利用することでシーンキーワードを抽出し、シーンのインデクシングを行う方式を提案する。特に、番組ジャンル毎のキーワードからなる定型句辞書と、番組固有のキーワードからなる動画固有辞書による多重インデクシング方式を提案する。これら二種類の意味レベルの異なる辞書により、指定シーンがトピック開始位置からズレる問題（検索位置ズ

レ)を解決する。また、キーワードの符号化による計算リソース削減方式を提案する。本研究成果の新規性は字幕情報を用いた多重インデクシング方式とシーンキーワードの符号化にある。また、その有用性は字幕情報利用で発生し得る検索位置ズレを補正できる点と、インデクシングにおける処理負荷を軽減し、必要な計算リソースを低減できる点にある。従来技術では、シーンのインデクシング処理に8MHzのCPU負荷と2MByteのメモリ量を必要とし、シーン検索精度を示すF値は65.8%程度である。提案方式により、シーン検索精度を示すF値を81.5%とすることができ、インデクシングに必要となるCPU負荷を1.835MHz、必要メモリ量を0.384MByteとすることができる。これにより、普及型の録画機能付きテレビに、CPUやメモリを変更或いは追加することなく、シーン検索機能を付加できることを示す。

第4章では、第2章で述べた要件を満たすテレビ番組推薦機能の実現技術として、ユーザの嗜好を的確に表現する必要最低限のテレビ操作と番組重要属性を用いた推薦スコア計算におけるリソース削減方式を提案する。また、推薦スコアと推薦アイテム数の関係を示す“Number-Score Curve (N-S Curve)”により推薦番組数を制御することで、テレビ番組推薦における適合率と再現率の最適化手法を提案する。これにより、最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とする。本研究の新規性はN-S Curveによる推薦番組数の決定方式とユーザの嗜好抽出における番組の重要属性及びテレビ操作の定義にある。また、その有用性は推薦における適合率と再現率のバランスを最適化できる点と、推薦における処理負荷を軽減し、必要となる計算リソースを低減できる点にある。従来技術では、300MHzのCPU、256MByteのメモリを搭載した普及型の録画機能付きテレビにおいて、一週間のテレビ番組(3000番組)の推薦処理に26秒を要する。また、推薦適合率と再現率は、それぞれ71%及び62%(F値=66.2%)である。これに対して提案方式では、同様の条件下で推薦処理を8秒で実施可能となる。また、テレビ番組推薦における適合率と再現率をそれぞれ82%及び78%(F値=79.9%)とすることができる。これにより

テレビ番組推薦機能における適合率と再現率のバランスを改善し、普及型の録画機能付きテレビに、CPUやメモリを変更或いは追加することなく、テレビ番組推薦機能を付加できることを示す。

第5章では、今後の展望として、近年様々な分野で活用が検討されているソーシャルメディアデータを活用したテレビ番組推薦及びシーン検索の可能性について述べる。ソーシャルメディアデータには、消費者のマインドセット（思考や価値観等）が多分に含まれており、このデータの活用に対する期待も高まっている。本研究では、ソーシャルメディアデータを活用した世の中のトレンドに基づくアイテムの推薦/検索の可能性を示す。特に、ユーザのアイテム選択履歴データとソーシャルメディアデータの融合分析による、世の中のトレンドの把握可能性を示す。これにより、今後のテレビ番組推薦及びシーン検索においても、これまでのようなユーザの嗜好に基づくだけでなく、世の中のトレンドにいち早く対応して推薦/検索することが可能となることを示す。

第6章では、本研究のまとめとして、本研究に対する結言と今後の課題について述べる。

第2章

テレビにおける付加機能実現とは

本研究のゴールは第1章で述べた大量テレビ番組の中からユーザが興味ある番組を見落とし無く、短時間で視聴可能とするためのシーン検索機能とテレビ番組推薦機能を録画機器に付加可能とすることである。録画機器としてはPC (Personal Computer) をはじめ多数の機器があるが、本研究では人々にとって身近である一方で、処理能力やユーザインターフェースに制限があるテレビを対象とする。中でも、近年普及が進んでいる録画機能付きのテレビを対象とする。テレビは子供からお年寄りまで、誰もが簡単に楽しめる家庭の中心的存在である。このため、コストや信頼性及び使い勝手に対する要求仕様は、PC等の他のデジタルメディア機器に比べて非常に厳しい。特に、テレビはコストを最適化した設計となっており、余剰能力を持たない。すなわち最大処理負荷時の処理能力に余裕がなく、他の処理をほとんど行えない設計となっている。一方で、今後も新しいテレビを市場に投入し、より良いテレビのコンセプトを消費者に提供するために新たな付加機能を実現する必要がある。しかし、付加機能を実現する場合には、価格を抑えるために設計変更を少なくし、CPUやメモリを変えないことも多い。また、老若男女を問わず誰でも使えるように、使い勝手の良い機能とする必要がある。さらに、ユーザに長く安心して使用してもらえるように信頼性のある機能とする必要がある。したがって、テレビに新たな付加機能を実現する際には、現状でも余

裕がないCPUやメモリを用いて安価で簡単かつ信頼性のある機能とすることが課題となる。これらの観点から、研究対象のシーン検索機能とテレビ番組推薦機能をCPUやメモリの変更或いは追加することなくテレビに付加可能とする。また、シーン検索においてはユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とする。テレビ番組の推薦においては最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とする。

2.1 はじめに

第1章で述べたとおり、近年、視聴可能なテレビ番組が大量化しており、これらの大量テレビ番組を見落とし無く、短時間で視聴したいというニーズが増えている。本研究では、これらの大量テレビ番組の中からユーザが好きなシーンのみを選択して視聴可能とするための「シーン検索機能」とユーザの好みに合ったテレビ番組を推薦/自動録画する「テレビ番組推薦機能」を録画機器に付加可能とする。

録画機器としては、HDR (Hard Disk Recoder)をはじめ、STB (Set-Top-Box) や携帯電話、PC等があるが、本研究では人々にとって身近である一方で機能実現に対する制約が厳しいテレビを対象とする。特に、テレビにはローエンドモデルからハイエンドモデルまで多くの種類があるが、本研究では普及型のエントリーモデルを対象とし、近年広く使われ始めた録画機能付きのテレビにシーン検索機能とテレビ番組推薦機能を付加可能とする。

以降では、本研究で対象としているテレビがどういうものか、その背景を踏まえて説明する。また、テレビの設計における要求事項を説明する。さらに、本研究においてシーン検索機能とテレビ番組推薦機能を実現する上での技術的な主眼点と問題意識を説明し、これらの機能に対する要求仕様を明確化する。そして、最後に、短時間視聴に関する従来の取り組みを説明し、本研究の有効性を示す。

2.2 テレビの概要

テレビは子供からお年寄りまで、誰もが簡単に楽しめる家庭の中心的存在である。これはテレビの普及率からもわかり、2014年3月時点で日本におけるテレビの普及率は96.5%とほぼすべての世帯に普及している[1]。パソコンや携帯電話も普及していると言われているが、それぞれ78.7%、93.2%であり[1]、テレビは極めて普及率が高く、人々にとって身近な製品となっている[1]。

テレビが市場に現れたのは1953年であるが、当時のテレビは消費者の収入に対して、かなり高価なものであった。しかし、現状では低コスト化が進み、この低コスト化が普及の大きな要因の一つとなっている。

テレビは家庭の中心的存在であり続けるために、低コスト化と並行して、これまでに多くの機能が搭載され、数々の進化を遂げてきている。すなわち、第1章でも述べたとおり、テレビは白黒からカラーへと変わり、やがてリモコンにより操作できるようになった。また、ポータブル化がなされてきた。現在、テレビはさらに進化を遂げ、デジタル放送により、高画質化及び多チャンネル化された番組を楽しむことができる。また、ネットワークインフラの発展と整備により、ネットワークを使用してオンデマンドでテレビ番組を視聴できるようになり、いわゆるネット動画の視聴も可能となっている。さらに、今後は3D(dimension)4K/8K解像度といった臨場感のある番組も一般化される可能性がある。

周辺装置としても、VTR(Video Tape Recorder)やHDR(Hard Disk Recorder)等が登場し、一般家庭でテレビ番組を容易に録画可能となった。近年ではこの録画機能もテレビに搭載され、さらにHDD(Hard Disk Drive)の大容量化やMPEG(Moving Picture Experts Group)等の動画圧縮技術の進化により、ユーザは好きなテレビ番組を簡単かつ大量に録画可能となっている。

現状、テレビには価格帯によってハイビジョン/フルハイビジョン及び画質

補正等の画質の違いや薄型/超薄型等のデザイン及びサイズ等の違いにより多数の種類が存在する．それだけではなく，先に述べた録画機能を一体化したものやスマートフォンと連携する機能も市場に投入されている．中でも録画機能付きテレビは低コスト化がなされ，2010年時点での市場シェアが3割を超える普及機となっている[1]．このように，テレビでは今後も低コストであり続けると共に，消費者に新しいコンセプトを提供するために新たな付加機能が実現されていくと考えられる．

2.3 設計における要求

前節で述べたとおり，テレビはその歴史的背景や身近さから，コストや信頼性及び使い勝手に対する要求仕様がPC等の他のデジタルメディア機器に比べて非常に厳しいものとなっている．特に，テレビで最も重要視されるポイントは画質や価格であるため[2]，新たな付加機能実現に費やせるコストは僅かである．したがって，本研究で対象とする普及型の録画機能付きテレビにおける付加機能の実現に対する要求仕様は厳しく，以下の通り計算リソースやユーザインターフェース及び信頼性を考慮する必要がある．

(1) 計算リソース

テレビに搭載されるCPUやメモリは非常に低コストなものとしており，普及型のテレビに搭載されているCPUやメモリの性能は現状のPCの1/10以下である．これは約15年前のPCと同等の性能であると同時に，最大処理負荷時の処理能力には余裕がなく，他の処理をほとんど行えないほどにコスト最適化された設計となっている．その一方で，録画時のデータ欠落や応答時間の低下がないように綿密に設計されている．近年では携帯電話やスマートフォンも普及してきているが，これらもノートPCと同等の性能となっており，テレビの処理能力はこれらの機器と比べて非常に制限が多い．このため，テレビに

における付加機能の実現要件は、これらの機器と比較しても著しく厳しく、より少ない計算リソースで付加機能を実現できるように、付加機能における処理負荷を軽減する必要がある。

(2) ユーザインターフェース

テレビを操作するためのユーザインターフェースは、原価わずか数百円のリモコンであり、ボタンによる操作が前提となっている。また、リモコンで離れた場所からテレビを操作することも多く、視覚上の表示解像度が低い。応答時間も考慮する必要がある。PCでは電源を投入してから数十秒待たないと使用できないが、歴史的な背景からかユーザはこれを許容しているか、あるいは許容せざるを得ないと考えている。しかし、テレビでは応答時間が長いことは許容されないと考えられる。つまり、スイッチを入れて数十秒待たなければ動作しないテレビやチャンネルを変更して数十秒も映らないテレビがユーザにとって身近な製品になるとは考えにくい。これらの制約からもテレビに付加機能を実現する際にはその機能に対するユーザインターフェースを良くするだけでなく、他の機能に対するユーザインターフェースに悪影響を与えないように設計する必要がある。

(3) 信頼性

信頼性に関しても歴史的な背景が関係しており、PCの場合には何らかの不具合で動作しなくなった場合、リセットしても許容せざるを得ないと考えられている風潮があるが、テレビでは許容されない可能性が高い。本研究で対象とするシーン検索機能やテレビ番組推薦機能に関しても、PCであれば検索結果や推薦結果が多少悪くても許容される可能性がある。しかし、テレビでは一度好ましくない結果が表示された場合に、その機能がユーザからのクレームの対象となる可能性や二度と使用されなくなる可能性もあるので十分な注意が必要となる。

これらの観点は、もちろん、コストを上げてCPUやメモリを追加あるいは高価なものに変更すれば解決できる可能性がある。しかし、テレビの世界は、販売価格が僅かに1000円高くなっただけで消費者から受け入れられなくなる世界である。価格に対してマスメディアから厳しく評価されることもある。あるいは原価コスト1000円増加に対して販売価格を据え置きにすると、数億円の利益を失う世界でもある。したがって、テレビに付加機能を実現する場合には、まずは可能な限りCPUやメモリを追加あるいは変更することなく、原価コストを上げずに実現可能とすることが重要となる。

一方で、既に述べたとおり、テレビはコストを最適化した設計としており、余剰能力を持たない。すなわち最大処理負荷時の処理能力には余裕がなく、他の処理をほとんど行えない設計となっている。したがって、テレビに付加機能を実現する際には、可能な限り現状でも余裕がないCPUやメモリを用いて付加対象の機能を実現することが課題となる。また、付加対象の機能を使い勝手良く、信頼性のあるものとすることが課題となる。

2.4 技術的な主眼点と問題意識

前節で述べた課題を解決するために、技術的には付加対象とする機能の処理負荷と精度を考慮することが必要となる。以下では、本研究で対象とするテレビ録画番組のシーン検索機能及びテレビ番組推薦機能をテレビに付加する際の処理負荷及び精度に対する技術的な主眼点と問題意識について述べる。

2.4.1 処理負荷

テレビ録画番組に対してシーン検索機能を実現する場合、当該録画番組のシーンを解析して、シーンのインデクシングを行う必要がある。これには、

テレビ番組の録画と同時にシーンのインデクシングを行う方法と、録画終了後にシーンのインデクシングを行う方法がある。しかし、録画終了後にインデクシングを行う場合、番組の録画が終了していてもインデクシングが終了するまでは、その番組のシーン検索を行えないため、ユーザインターフェース上好ましくない。よって、本研究では録画中にリアルタイムにインデクシングを行うことを想定する。先に述べたとおり、テレビにおいて付加機能を実現するには、価格を抑えるために設計変更を少なくし、CPUやメモリを変えないことが要求される場合も多い。一方で、通常、テレビはコスト最適化された設計となっており、録画状態での最大処理負荷時の計算リソースには余裕がなく、CPU利用率やメモリ使用率が100%に達すると録画データの欠落やユーザインターフェースにおける極端な応答速度の低下を招くことになる。このため、どのテレビであっても、設計上はCPU利用率やメモリ使用率が100%とならないようにマージンを持っている。

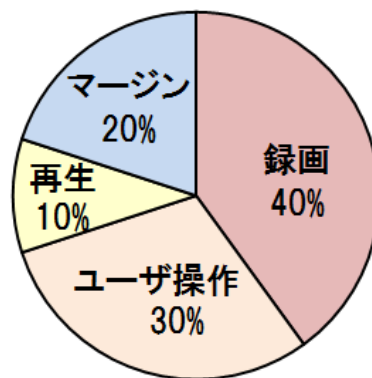


図2-1 録画時の計算リソース使用率に関する設計思想

図2-1は300MHzで動作するCPUと256MByteのメモリを搭載した普及型の録画機能付きテレビでの実測結果に基づく、録画状態での計算リソース使用率に関する設計思想を示している。ただし、これは定常状態での設計思想であり、通常は入出力コンフリクトやユーザからの想定外の操作があるため、実際にはCPU利用率が想定を超える場合がある。

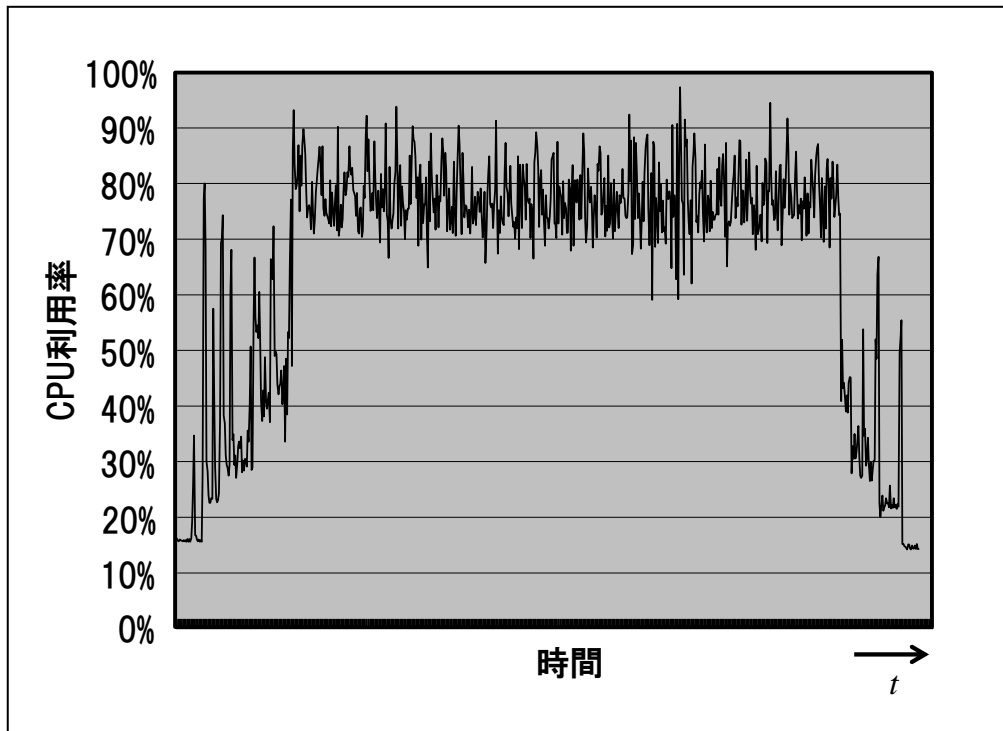


図2-2 録画時のCPU利用率

図2-2は前述した設計思想に基づいて実現した，実際の普及型の録画機能付きテレビにおける録画状態でのCPU利用率の測定結果を示したものである．図2-2に示す通り，最大処理負荷時のCPU利用率が80%となるように設計されていても，実際には97%に達する場合があった．また，メモリ使用率も99%を超える場合があった．先に述べたとおり，これらのいずれかでも100%に達すると，録画データの欠落やユーザインタフェースにおける極端な応答速度の低下を招くことになる．このため，普及型のテレビでは，CPU及びメモリは全体の1%ほどしか使えない．そこで，本研究では，シーン検索機能をCPU利用率1%以下，使用メモリ量0.2%以下で実現可能とすることを設計要求とした．特に，普及型の録画機能付きテレビを対象として，300MHzで動作するCPUと256MByteのメモリを搭載した録画機能付きテレビにおいて，CPU負荷を3MHz以下，使用メモリ量を0.5MByte以下で実現可能とすることを設計要求とした．これを実

現しなければ、最大処理負荷時にCPU利用率及びメモリ使用率が100%に達する可能性があるため、CPUやメモリの変更或いは追加が必要となり結果的に普及型のテレビにシーン検索機能を付加することが不可能となりうる。これらはある一つの録画機能付きテレビを対象として示しており、データとしても同テレビで測定した結果を示しているが、考え方は基本的かつ普遍的であると考えられる。すなわち、前述したとおり、普及型のテレビが人々にとって身近な存在である以上、コスト最適化した設計とし続ける必要があり、計算リソースの使用量も厳しい設計に基づいて決められていることは変わらない。このため、どの製品や機種および時代についても同様の考え方が用いられると考えられる。

また、テレビ番組の推薦では、夜中等のテレビが動作していない待機時に推薦すべき番組を計算することも可能である。しかし、待機中にユーザがテレビを起動する可能性もあるため、待機中に多くの処理を行うことは望ましくない。また、待機電力を抑えるため、テレビを使用していない時には主電源を切るユーザもいる。さらに、放送される番組や終了する番組は時々刻々と変化することを考慮し、常に最新のテレビ番組が推薦できるように配慮する必要がある。そのため、本研究では、ユーザからリクエストがあった際に推薦計算を行う仕様とした。この場合、ユーザからのリクエストに対して推薦番組をリアルタイムに計算する必要があり、可能な限り高速に推薦処理を実施する必要がある。具体的には、ユーザがあるリクエストをテレビに対して行った際に、ユーザが待つことができる時間は10秒程度であると考え、推薦処理も10秒以下で完了することを設計要求とした。これを実現しなければ、テレビ番組推薦機能の使い勝手が悪いあるいは他の機能の使い勝手に悪影響を及ぼすため、やはりCPUやメモリの変更或いは追加が必要となる。結果的に原価コストが上がってしまうため、普及型のテレビにテレビ番組推薦機能を付加することが不可能となりうる。

2.4.2 精度

シーン検索機能では、ユーザから指定されたシーンが精度よく検出される必要がある。通常、シーン検索ではあるキーワードや顔画像を指定した場合に、そのキーワードや顔が出現するシーンを検出する。このため、指定されたシーンに関するトピックの開始位置からはズレが生じ、指定シーンのトピックの最初から再生を行うことができないため、ユーザにとっては精度が良い検索とは言えない。具体的には、例えばユーザが出演者名を指定した場合に、その出演者が出ているトピックの途中から再生しても、ユーザにとっては、その前後関係がわからず、番組を楽しむことはできない。そこで、本研究では、シーン検索においてユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とするシーン検索機能を実現する。これを実現しなければ、シーン推薦機能の使い勝手が悪いあるいは信頼性のない機能と判断され、ユーザに使用されない機能となりうるため、信頼性が強く要求されるテレビにシーン検索機能を付加するという判断にはなりえない可能性がある。

テレビ番組の推薦では、推薦番組数が多い場合には、ユーザは多くの選択肢を得ることが可能となるが、好みではない番組も推薦される可能性があり、所望の番組を選択することが困難となる。一方で、推薦番組数が少ない場合には、ユーザは好みの番組を選択することが容易となるが、選択肢が少なく、所望の番組が推薦されない可能性がある(表2-1)。すなわち、これらはトレードオフであり、推薦においてはこれらのバランスが重要となる。特に、テレビでは、視覚上の解像度が低い画面に表示された推薦番組をリモコンによって簡単に選択できるようにする必要がある。また、番組自動録画のためのハードディスク容量や記憶容量も限られており、これらを浪費することなく、最適な番組を自動録画する必要がある。そこで、本研究では、最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とするテレビ番組推薦機能

実現する．これを実現しなければ，テレビ番組推薦機能の使い勝手が悪いあるいは信頼性のない機能と判断され，ユーザに使用されない機能となりうるため，やはりテレビにテレビ番組推薦機能を付加するという判断にはなりえない可能性がある．

表2-1 推薦アイテム数に関するトレードオフ

推薦番組数	メリット	デメリット
多い	選択肢が多い	選択困難
少ない	選択容易	選択肢が少ない

2.5 処理負荷と精度に関する位置づけ

前節で述べた主眼点と問題意識に基づき，本研究では可能な限り処理負荷を低減し，必要計算リソースを軽減した上で，精度の良いシーン検索機能及びテレビ番組推薦機能の実現をターゲットとする．もちろん，図2-3に示すとおり，計算リソースに制限がなく，処理負荷が大きくても良いのであれば，

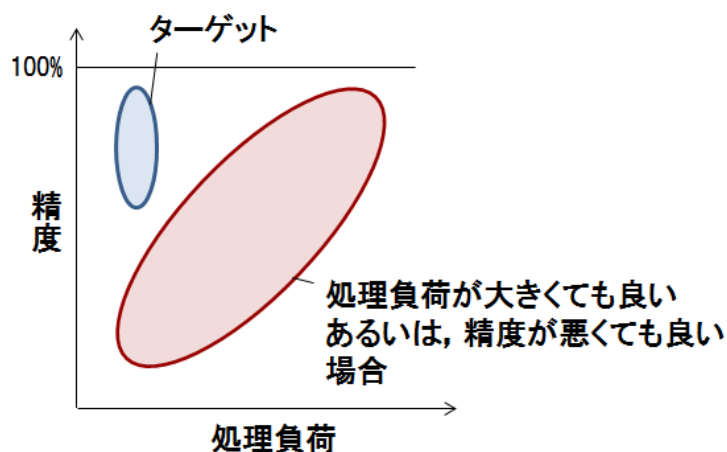


図2-3 処理負荷と精度の位置づけ

精度の良いシーン検索機能及びテレビ番組推薦機能を実現できる可能性がある。逆に、精度が悪くても良いのであれば、処理負荷を少なくし、必要計算リソースを低減することが可能である。しかしながら、テレビへの付加機能の実現では、前述したとおり必要計算リソースの低減と精度の良さを両立する必要がある。本研究ではこれらを両立する技術の実現をターゲットとする。

2.6 短時間視聴に関する従来の取り組み

著者らは本研究にてシーン検索機能を実現する前に、録画テレビ番組のシーン推薦機能を実現し、PCに搭載している[3][4]。実現したシーン推薦機能は、テレビ録画番組における映像と音声を解析することで観どころと思われるシーンを抽出し、これらのシーンを再生することで、観どころシーンのダイジェスト視聴を可能とする。具体的には、映像解析ではテロップ検出やシーンチェンジ検出及び特定絵柄検出等[5]を行い、音声解析では盛り上がり検出や楽曲検出等[6]を行うことで、録画テレビ番組における盛り上がりシーン等の観どころシーンを抽出して再生する。これにより、例えば、2時間のサッカー番組を10分で楽しむことができるようになる。

本機能では、録画テレビ番組の映像と音声を解析するため、非常に多くの計算リソースを必要とする。特に、テレビ番組はMPEG-2 (Moving Picture Experts Group-2) で圧縮されており、このデコードには多くの計算リソースを使用する。このため、本実現ではダウンデコード方式による低処理負荷化を提案した。提案したダウンデコードでは、MPEGストリームに含まれるDCT (Discrete Cosine Transform) 係数の一部を用いて高精細画像を標準精細画像にダウンデコードすることで、デコードに要するCPU負荷とメモリ量を削減できる[7]。また、高精度でのシーン抽出を可能とするために、ドリフトリダクションフィルタと呼ぶ水平／垂直ともに4 タップのフィルタにより、周辺の画素から参照画素をリアルタイムに補間し、高画質画像処理を行うことが

できる[8]. これにより, 平均再生時間を全体の20%とした場合の観どころシーン抽出適合率を67%, 再現率を74%とできる. しかし, 精度に関する議論となる前に, 使用計算リソースが問題となった. 本機能では, CPU負荷285MHz, メモリ使用量は32MByteとなり, 300MHzで動作するCPUと256MByteのメモリしか持たない普及型のテレビでは, CPU及びメモリを追加あるいは高性能なものに変更する必要があった. 実際には, 本機能を日立製のメディアプロセッサBroadGear[9][10]において, これらの機能を実現することを提案したが, 原価コストが上がってしまうため, CPUの追加は困難であり, 結果的にテレビにシーン推薦機能を付加することはできなかった.

シーン推薦機能に関しては, 音声情報のみを用いて, シーン内の発話時間長の遷移により推薦シーンを決定することで実現する技術[11]や, ステレオ音声からL(Left)-R(Right)の左右差分信号を用いて盛り上がりのシーンを検出することで実現する技術も提案されている[12]. また, 映像のみを用いたシーンの遷移確率に基づくイベント検出による野球番組のインデクシング方法も提案されている[13]. さらに, 動画の圧縮状態でマクロブロックタイプやDCT係数の特徴を解析することで野球番組におけるゴールシーン等を検出し, これらのシーンを再生することでダイジェスト視聴を可能とする技術も提案されている[14]. これらの技術により処理負荷を低減し, 使用計算リソースを抑えてシーン推薦機能を実現することができる. しかしながら, これらの技術では具体的な処理負荷及び必要計算リソースは明確にされていないものの, 映像や音声データといった大量のデータを処理する必要があるため, 少なくとも数十MHzのCPU負荷と数MByteのメモリを消費してしまうと考えられる. このため, やはりCPUやメモリを追加あるいは高性能なものに変更する必要があり, 普及型のテレビにこれらの機能を付加することは困難であると考えられる.

本研究では, 大量テレビ番組の中からユーザが興味ある番組を見落とし無く短時間で視聴可能とするために, 以下の機能を録画機器において実現する.

第2章 テレビにおける付加機能実現とは

- ・ユーザが好みのシーンを検索して視聴可能とする「シーン検索機能」
- ・ユーザの好みに合った番組を推薦/自動録画可能とする「テレビ番組推薦機能」

特に、処理能力やユーザインターフェースに制限がある普及型の録画機能付きのテレビにこれらの機能を付加可能とするため以下を実現する．

- ・普及型の録画機能付きのテレビにおいて、CPUやメモリの変更あるいは追加することなくシーン検索機能及びテレビ番組推薦機能を付加可能とする．特に、普及型のテレビではCPU及びメモリは全体の1%ほどしか使えないことを考慮して、具体的に300MHzのCPU、256MByteのメモリを持つ普及型の録画機能付きテレビにおいて、シーン検索処理をCPU負荷3MHz、使用メモリ量0.5MByte以下で実現可能とする．また、テレビ番組の推薦では、ユーザからのリクエストに対して、ユーザが待つことができると考えられる10秒以下で推薦処理を完了可能とする．
- ・シーン検索において、ユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とする．また、テレビ番組の推薦において、最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とする．

これにより、消費者にとって身近なテレビにおいて、好きなテレビ番組を見落としなく素早く選択して、そのシーンを短時間で視聴可能となる．そして、近年多忙化する生活者が低価格なテレビを用いて、短時間で大量のテレビ番組を楽しむことが可能となる．

新たな付加機能実現の必要性は、テレビに限らず、HDR(Hard Disk Recorder)やSTB(Set-Top-Box)でも同様である．また、コンシューマ製品に限らず、サーバ等のプロバイダ製品でも同様であり、高精度化や必要計算リソース軽減に向けた処理負荷の低減は、ユーザインターフェースの向上や信頼性及び高速処理につながる．このため、本研究で得られる成果はテレビに限らず多くの製品に対して有効であると考えられる．

2.7 第2章まとめ

本章では、本研究で実現する「シーン検索機能」と「テレビ番組推薦機能」の搭載対象であるテレビの概要を説明した。また、テレビの設計における要求事項を説明した。さらに、シーン検索機能とテレビ番組推薦機能をテレビに付加可能とする上での技術的な主眼点と問題意識を説明し、実現すべきシーン検索機能とテレビ番組推薦機能の要求仕様を明確化した。

テレビは一般家庭に普及して以来、子供からお年寄りまで、誰もが簡単に楽しめる家庭の中心的存在である。その歴史的背景や身近さから、コストや信頼性及び使い勝手に対する要求仕様は、PC等の他のデジタルメディア機器に比べて非常に厳しいものとなっている。特に、テレビはコストを最適化した設計になっており、余剰能力を持たないため、最大処理負荷時の処理能力に余裕がなく、他の処理はほとんど行えない設計となっている。一方で、今後も新しいテレビを市場に投入し、より良いコンセプトを消費者に提供するために新たな付加機能を実現する必要がある。しかし、付加機能を実現する場合には、価格を抑えるために設計変更を少なくし、CPUやメモリを変えないことも多い。また、テレビに付加される機能は、ユーザに長く安心して使用してもらえる機能とするために、使い勝手の良い信頼性のある機能とする必要がある。したがって、本研究では、テレビにおいてCPUやメモリを変更或いは追加することなく、シーン検索機能及びテレビ番組推薦機能を付加可能とする。また、シーン検索において、ユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とする。テレビ番組の推薦では、最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とする。これにより、消費者にとって身近なテレビにおいて、好きなテレビ番組を素早く選択して、そのシーンを短時間で視聴可能とする。そして、近年多忙化する生活者が低価格なテレビを用いて、短時間で大量のテレビ番組を楽しむことを

可能とする。

新たな付加機能実現の必要性は、テレビに限らず、HDR(Hard Disk Recoder)やSTB(Set-Top-Box)でも同じである。また、コンシューマ製品に限らず、サーバ等のプロバイダ製品でも同じであり、高精度化や必要計算リソース軽減に向けた処理負荷の低減は、ユーザインタフェースの向上や信頼性及び高速処理につながる。このため、本研究で得られる成果はテレビに限らず多くの製品に対して有効であると考えられる。

第3章

シーン検索技術

本章では、第2章で述べたとおり、処理能力及びユーザインターフェースに制限があるテレビにおいて、近年大量化するテレビ番組の中から好きなシーンのみを選択して視聴可能とするための「シーン検索機能」を提案する。特に、テレビ録画コンテンツに含まれる字幕情報を用いた低処理負荷/高精度シーンインデクシング方式と使い勝手の良いシーン検索機能を提案する。字幕情報は基本的にはテレビ番組内でしゃべられている音声そのままテキスト化されている。このため、キーワード（例えば人物名等）は必ずしもトピックの最初だけに出てくるとは限らず、キーワードを選択してもトピックの最初から視聴できないという「検索位置ズレ」の問題が発生する。そこで、提案方式では、テレビ番組のジャンル毎に共通なキーワードからなる定型句辞書と、番組に固有なキーワードからなる動画固有辞書による多重インデクシングを行う。これら二種類の意味レベルの異なる番組キーワード辞書を活用することで、字幕情報利用で発生し得る「検索位置ズレ」の問題を解決する。また、シーンに対して複数の意味を持たせ、大まかなシーンのザッピング視聴と詳細なシーンの選択視聴を実現する。さらに、低処理負荷での実現に向けて、キーワードを符号化することによってシーンの効率的な表現を提案する。

3.1 はじめに

第1章で述べたとおり、近年、視聴可能なテレビ番組が大量化している。特に、録画機能が一般化し、多くのテレビ番組を録画可能となり、これらの録画番組を短時間に視聴したいというニーズが増えている。そこで、本研究では、普及型の録画機能付きテレビにおいて、録画されたテレビ番組の中から好きなシーンのみを選択して視聴可能とするための「シーン検索機能」を提案する。特に、300MHzのCPU、256MByteのメモリを持つ普及型の録画機能付きテレビにおいて、CPUやメモリを追加あるいは変更せず、原価コストを上げることなく、「シーン検索機能」を付加可能とする。通常、テレビは最大処理負荷時の計算リソースに余裕がなく、他の処理をほとんど行えないほどにコスト最適化された設計となっている。このため、第2章で述べたとおり、CPUやメモリを増加あるいは変更することなく、これらの機能をテレビに付加可能とするためには、CPU負荷3MHz以下、使用メモリ量0.5MByte以下でシーン検索に必要なインデクシング処理を実現可能とする必要がある。また、シーン検索ではユーザが指定したシーンに関するトピックの最初から正確に視聴できることが要求される。そこで、本研究では、原価コストを上げることなく、ユーザが指定したシーンのトピックを正確に視聴可能とすることを目標とし、図3-1に示すシーン検索機能を実現する。

《目標》

- (1) CPU負荷3MHz, メモリ量0.5MByteで機能を実現可能とすること
- (2) ユーザが指定したシーンのトピックを正確に視聴可能とすること

図3-1 シーン検索機能の実現目標

3.2 関連技術

動画を構成するメディアとしては映像，音声，テキストがある．テレビ番組中のシーン検索機能を実現するための技術としても，これらのテレビ番組中の動画における映像，音声，字幕情報（テキスト）あるいはこれらを組合わせた解析による方法がある（図3-2）．これらの方法では，動画中の特徴的なシーンや人物のシーンを検出して，それらのシーンに対するキーワードを生成することでインデックス（或いはメタデータとも言う）を生成し，動画中のシーン検索を可能とする．以下ではこれらのシーン検索機能を実現するための技術について詳述する．

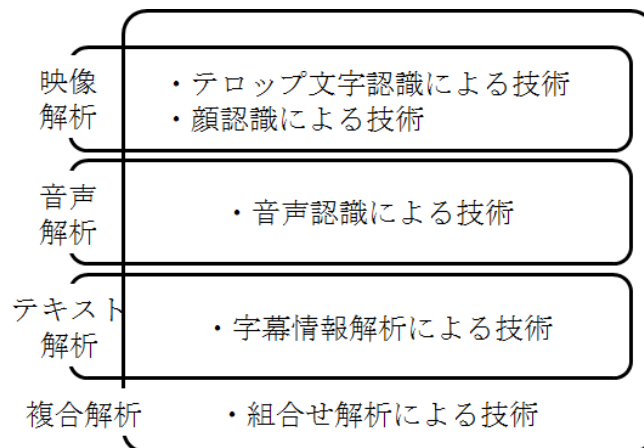


図3-2 シーン検索機能を実現する技術

映像解析による技術としては，映像に組込まれたテロップ文字列を認識あるいは映像中に映っている顔を認識することで，動画中のシーンに対するキーワードを生成し，シーンのインデクシングを行う方法が知られている[15][16]．

テロップ内の文字列を認識する技術としては，テロップ内の文字の色やサイズ等の制限を特徴として利用し，文字列の切れ目を補正することでテロッ

プ領域を検出して、文字列を認識する技術が提案されている[17]．これにより、動画中のテロップ文字列を正確に認識し、シーンのインデクシングが可能となる．また、複数フレームにまたがって、文字列の位置を評価することで、多様なフォントサイズ及びスタイルを持つ文字列の認識を可能とする技術も提案されている[18][19]．このように複数フレームを利用することで、複雑な背景映像であった場合でもテロップ内の文字列を認識可能となる．このため、より多くのテロップ文字列を認識し、より多くのシーンにインデックスを付けることが可能となる．テロップ文字列の認識においては、テロップ領域の検出が重要であるが、これを実現する技術としても、映像中のコントラストや動きを分析することでテロップ領域を検出する技術が提案されている[20]．これによって、動きのあるテロップを検出し、これらで表示された文字列を用いてシーンのインデクシングを行うことが可能となる．また、映像における背景とテロップ領域の色及び輝度の差によってテロップ領域のエッジを検出し、背景と二値化することでテロップ領域を抽出する方法も提案されている[21][22]．これにより、複数フレームを用いることなく単一フレームで背景ノイズを除去した高速なテロップ領域の抽出が可能となる．さらに、最近では、文字の回転や視点の変化に対応したテロップ文字列の認識手法も提案されている[23]．これにより、映像中の看板等の自然画における文字を認識し、これをインデクシングに活用することも可能となる．これらの技術では80%以上の認識率でテロップ文字列を認識できる[23]．また、テロップ（例えば人物名等）が表示されるタイミングはトピックの変化点で表示されることが多いため、トピックの開始位置に対して、キーワードによる検索位置ズレの少ないシーン検索が可能である．しかしながら、これらのテロップを活用する手法では、1秒あたりに1画像(Frame)を処理するとしても、テロップの抽出だけで200MHzのCPU負荷を消費する[21]．また、現在のデジタルテレビ放送において主流となっている高精細映像を処理する場合、少なくとも4MByte以上のメモリを必要とする．加えて、1キーワード当たり14文字

(符号値4バイトを合わせて32Byte)のシーンキーワード辞書を8万語分持つとすると、さらに2.56MByte (8万語×32Byte)、合計6MByte以上のメモリが必要となる。このため、これらの技術では、本研究で目標としているCPU負荷3MHz以下、メモリ0.5MByte以下での実現は困難である。

顔を認識する技術としては、映像フレームの時間的連続性を利用し、同一人物をクラスタリングして顔認識する技術が提案されている[24]。これにより、動画に特有な撮影環境の違いを吸収し、映像中の照明や顔向きに影響されずに人物シーンのインデクシングを行うことが可能となる。また、顔の表情表出時における眉や目の端点等の顔特徴点を考慮して、顔を認識する技術も提案されている[25]。これにより、顔の表情に影響されることなく人物シーンのインデクシングを行うことが可能となる。顔認識技術は主に監視映像における人物検索等にも使われているが、テレビ番組における好みの人物を検索する場合にも有効であると考えられ、実際に音楽番組のシーンのインデクシングを行う技術も提案されている[26]。この技術を用いることにより、音楽番組におけるボーカリストの登場シーンに対してインデクシングを行うことが可能となる。認識までは行わず、テレビ番組内の顔検出のみを実施して、顔画像をサムネールとして一覧表示する製品も紹介されている[27]。この製品ではシーンのキーワードをユーザに提示することはできないが、ユーザは顔画像のサムネールを確認して、所望の人物が出現しているシーンを視聴可能となる。これらの技術では顔画像による人物の認識率は約60%である[25]。また、顔画像は多くのシーンで表示されるため、トピックの変化点に正確にインデクシングすることは困難であり、認識された人物名をキーワードとしてインデクシングしたとしても、その人物が現れるトピックの開始位置とはズレた位置にインデクシングしてしまい、検索位置ズレが生じやすい。このため、シーンの検索精度はテロップ認識による技術と比較して低くなる傾向がある。また、これらの顔認識による手法では、1秒あたりに1画像(Frame)を処理するとしても、顔画像の抽出だけでも多くの処理を必要とす

る．従来から顔認識における処理負荷に対する検討が行われてこなかったとして，分岐限定法や局所特微量の活用による効率的な顔認識方法も提案されている[28][29]．本技術により顔認識を高速に実現可能となる．しかし，それでも1Frameの処理には240MHz以上のCPU負荷を消費する．また，高精細映像を処理する場合，4MByte以上のメモリが必要となり，これにシーンキーワード辞書も加えるとテロップ認識と同様に6MByte以上のメモリが必要となる．このため，これらの技術では，本研究で目標としているCPU負荷3MHz以下，メモリ0.5MByte以下での実現は困難である．

音声解析による技術としては音声認識技術を応用した技術が提案されている[30]．これにより，動画中でしゃべられている音声に基づいて，動画におけるシーンのキーワードを抽出し，シーンのインデクシングを行うことが可能となる．音声認識技術そのものは古くから研究されているが，その処理負荷は多く，1秒の音声に対して300MHz程度のCPU負荷を消費する[31]．これに対して，言語モデルの粗密探索や探索閾値の導入により，音声認識における処理負荷を軽減する技術が提案されている[32]．これにより，126.5MHzのCPU負荷で音声認識を行うことが可能となる．また，近似GMM(Gaussian Mixture Model)演算の導入により処理負荷を削減する技術が提案されており[33]，これにより83.3MHzのCPU負荷で音声認識処理が可能となる．さらに，単語探索処理におけるtri-gramの適用と処理の並列化方式が提案されており[34]，これにより62.5MHzのCPU負荷で音声認識処理が可能となる．このように，音声認識処理における数々の低処理負荷化及び高速化が実現されているものの，目的としているCPU負荷3MHz以下での実現には至っていない．一方で，使用メモリとしては，音声認識に必要なメモリ自体は0.9MByte程度[32]と映像解析による技術に対して低減できる．しかし，シーンキーワード辞書2.56MByteを加えると3MByte以上のメモリが必要となり，やはり目的としているCPU負荷3MHz以下，メモリ0.5MByte以下での実現は困難である．また，音声認識の認識率は90%以上[32]とテロップ認識や顔認識よりも高い一方で，しゃ

べられるキーワード（例えば人物名等）は必ずしもトピックの最初だけに出てくるとは限らない．このため，トピックの開始位置に正確にインデクシングすることは困難であり，トピックの開始位置に対する検索位置ズレを生じ，シーン検索精度としては劣化してしまう．

テキスト解析による技術としては，EPG（Electronic Program Guide）における番組紹介文と字幕情報の共通単語を検出することで，シーンのインデクシングを行う方法が提案されている[35]．これにより，番組紹介文で出現している単語をキーワードとしてシーンのインデクシングを行うことが可能となる．また，字幕情報における言い回しの特徴に基づいて映像シーンに登場している主要な被写体を推定する方法[36][37]や反復語を捉えることで，シーンの特徴キーワードを抽出する方法も提案されている[38]．これらにより，特定のシーンにキーワードを付与し，シーンのインデクシングを行うことが可能となる．製品としても，字幕情報からタレント名を検出して，シーンのインデクシングを行う機能が登場している[39]．これにより，ユーザが入力した好みのタレントが登場するシーンをリストアップし，シーンの頭だし再生を行うことが可能となる．これらの字幕情報を用いた技術に関しては，いずれも処理負荷及び必要計算リソースに関しては明確にされていないが，映像や音声を解析する場合に比べて処理するデータ量が少なく，映像や音声及び文字の認識も行う必要がない．このため，映像や音声を解析する場合に比べて少ないCPU負荷での機能実現が可能である．しかしながら，これらの技術では，いずれも字幕情報の中から不定キーワードを抽出する必要があり，通常は形態素解析[40]等の自然言語処理を行う必要がある．キーワード抽出を効率的に行うためのダブル配列法[41]も提案されている．これによりインデクシングに必要となるキーワード抽出を高速に実施することが可能となる．しかし，それでも，字幕情報が1秒間に30文字，5個程度の単語を含むと仮定した場合に8MHz程度のCPU負荷を消費してしまう．また，処理に必要なメモリ量に関しても，映像や音声を解析する場合に比べて小さくはできるもの

の、シーンインデクシング用のキーワード辞書は必要となるため、上記同様8万語程度のキーワード辞書を用いる場合には2MByte以上のメモリが必要となる。このため、これらの技術でも目的としているCPU負荷3MHz以下、メモリ0.5MByte以下での実現は困難である。シーン検索精度という観点では、字幕情報を用いた解析では認識処理が必要ないため高精度にシーンのキーワードを生成できる。しかし、字幕情報は基本的には動画内でしゃべられている音声そのままテキスト化されているため、音声認識を行う場合と同様に、しゃべられるキーワード（例えば人物名等）は必ずしもトピックの最初だけに出てくるとは限らない。このため、音声認識による技術と同様に、トピックの開始位置に対する検索位置ズレを生じ、シーン検索精度としては劣化してしまう。

映像と音声及び字幕情報の解析を2つ以上組み合わせた複合解析による技術も提案されている。映像をベースとした複合方式としては、テロップ表示の変化とカメラワークの特徴に基づいて、野球番組における打席結果の推定方法が提案されている[42]。これにより、野球番組における得点シーン等の特定イベントシーンにインデクシングを行うことが可能となる。また、ニュース番組特化して、シーン中の顔の個数に基づいて映像を分類し、その分類に応じたシーンの見出しキーワードを字幕情報から抽出する技術が提案されている[43]。本技術によりニュース番組における人間に関するキーワードを生成し、これらのシーンのインデクシングを行うことが可能となる。同様に、ニュース番組に特化した技術として、顔抽出とテロップ認識を活用したAnchor ShotとReportシーンを分割する技術も提案されている[44]。本技術により、ニュース番組におけるトピックの変化ポイントに対してインデクシングを行うことが可能となる。また、顔抽出によって抽出した顔の人物名を字幕応用及びテロップ認識によって判別する技術も提案されている[45]。これにより、映像中の人物が現れているシーンに対して、人物名をキーワードとしてインデクシングを行うことが可能となる。音声をベースとした複合方式

としては、話者の出現確率に基づいて、ドラマ、アニメ、バラエティ番組におけるシーンの話者を推定する方式が提案されている[46][47]．これによって、話者毎にシーンのインデクシングを行うことが可能となる．また、音楽検出及び音声検出に、カメラワークの解析とテロップ検出を組み合わせ、シーンメタデータの作成を支援するシステムも提案されている[48]．本システムは人手によるシーンインデクシングを支援するものであるが、シーンのタイトル及び概要文を基にキーワードを生成してシーンのインデクシングを行うことができる．同様に人手によるシーンメタデータの生成を支援するシステムとしては、テロップ認識及び音声認識結果と字幕情報を自然言語処理することで、シーンインデクシングを支援する技術が提案されている[49]．これらの技術では、それぞれの単一方式におけるシーン検出の不完全性を補い、インデクシングの誤りを補正することが可能となる．したがって、トピックの開始位置に対する検索位置ズレの少ないシーン検索を実現することが可能である．一方で、これらの技術に関しては、処理負荷や必要計算リソースに関する明確な記載はない．しかし、先に述べたとおり、単一方式によるインデクシングも本研究で目標としているCPU負荷及びメモリ量での実現が困難であるため、これらの技術を組み合わせた技術で、目的としているCPU負荷3MHz以下、メモリ0.5MByte以下でシーン検索機能を実現することは困難であると言える．

表3-1 シーン検索関連技術の本研究目標達成に対する近さ

分類	処理負荷	検索位置ズレ
テロップ文字認識による技術	×	○
顔認識による技術	×	×
音声認識による技術	×	△
字幕情報解析による技術	△	△
組合せ解析による技術	××	◎

以上をまとめると各関連技術の本研究目標達成に対する近さは表3-1の通りとなる。

表3-1に示す通り、関連技術のいずれにおいても、本研究で目標としている「CPU負荷3MHz以下、使用メモリ量0.5MByte以下で、ユーザが指定したシーンに関するトピックの最初から正確に視聴可能とする」シーン検索機能を実現することは困難である。特に、必要計算リソースについては、いずれの関連技術でもCPU負荷3MHz以下、使用メモリ量0.5MByte以下でのシーン検索機能の実現が不可能である。このため、前節で述べた通り、普及型のテレビでこれらの機能を実現する場合には、CPUやメモリ等を増加或いは変更する等、ハードウェア構成を変更する必要がある。原価コストを上げる必要がある。そこで、本研究では、関連技術の中でも最も処理負荷を軽減し、必要計算リソースを低減することができる字幕情報を活用して、更なる低処理負荷化による低計算リソース化を実現する。また、字幕情報の利用においてトピックの開始位置に正確にインデクシングすることができない（検索位置ズレ）問題を解決する。これらにより、関連技術では両立できていない、本研究で目標としている「CPU負荷3MHz以下、メモリ0.5MByte以下でのシーンのインデクシングの実現と、ユーザが指定したシーンのトピックを正確に視聴可能とする」シーン検索機能を実現し、普及型の録画機能付きテレビにおいて原価コストを上げることなくシーン検索機能を付加可能とする。

3.3 提案方式

前述の目標を実現するため、以下のステップによるシーンのインデクシング方式を提案する(図3-3)。

ステップ1：番組のTS(Transport Stream)から字幕ES(Elementary Stream)を抽出し、字幕情報(サブタイトル情報)を取得する。

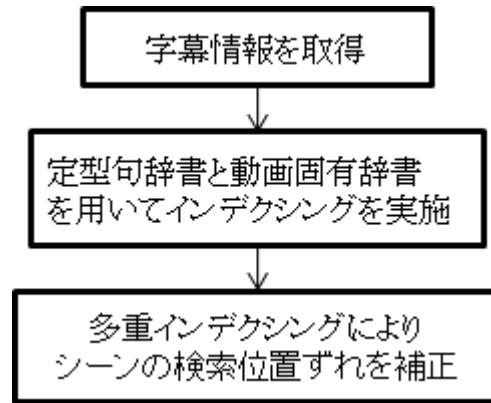


図3-3 提案方式における処理の流れ

ステップ2：字幕情報と定型句辞書及び動画固有辞書の2つの辞書を用いて、番組シーンのインデクシングを行う。

ステップ3：定型句辞書と動画固有辞書によるインデクシング結果を組合わせた多重インデクシングにより、シーンの検索位置ズレを補正する。

以下では、字幕情報と、本方式を実現する定型句辞書、動画固有辞書、多重インデクシング法について述べる。

3.3.1 字幕情報

字幕情報は、図3-4に示す通り、テレビ番組のTS(Transport Stream)中にES(Elementary Stream)として、映像や音声とは別に含まれているテキストデータである。テキストデータは映像データや音声データに比べてデータ量が少ない。それゆえ、字幕情報の処理による機能実現では、映像情報や音声情

報の処理による機能実現に比べ、処理負荷を軽減し、必要な計算リソースを軽減することができる。

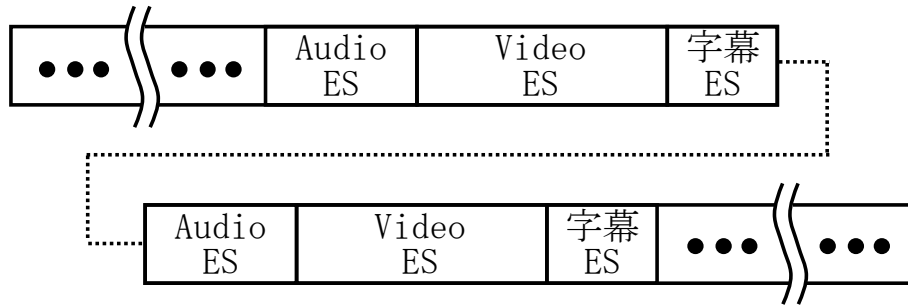


図3-4 Transport Stream

3.3.2 定型句辞書

定型句辞書は、テレビ番組のジャンルに対して一般的かつ共通なキーワードを集めた辞書であり、ジャンル毎に「続いては、」のように話題の変化点やイベントの開始位置を示すキーワードを集めた辞書である。キーワードは、著者らが実際にニュース、音楽、バラエティ、サッカー、野球の各番組をそれぞれ20番組、合計100番組を視聴することで調査して設定した。例えば、ニュース番組に対しては、トピックの始まりで表示される「続いては、」や「さて、」及び「ところで、」等を設定した。また、音楽番組に対しては、曲の始まりで表示される「曲です。」や「お聴きください。」等を設定した。このように、テレビ番組のジャンル毎に話題の変化点やイベントの開始位置で表示される数十個のキーワードを著者ら自身が見つけ出して設定した。

本辞書の音楽番組での利用方法を例示する。表3-2に音楽番組用の定型句辞書を示す。字幕情報内に検出キーワードの文字列が現れた時刻に、提示キーワードを用いてインデクシングする。これにより、曲の開始位置にインデックスを付けることができ、ユーザは楽曲シーンを検索可能となる。同様に、

表3-2 定型句辞書の例

#	検出キーワード	提示キーワード
1	曲です.	曲
2	お聴きください.	曲
3	聞いてください.	曲
4	見てください.	曲
5	どうぞ.	曲
...

本辞書により各種の番組ジャンルに対して一般的なトピックやイベントの開始位置を検索可能となる.

本定型句辞書は一般的な意味レベルのキーワードを持つ. このため, 本辞書によるインデックスは多数生成されるという短所がある. 一方, 生成されるインデックスの位置は, 著者ら自身が見つけ出した変化点やイベントの開始位置で表示されるキーワードに基づいて設定されるため, 正確である場合が多いという長所がある.

3.3.3 動画固有辞書

動画固有辞書は, 番組毎に固有かつ特徴的なキーワードを集めた辞書であり, 出演者名や曲名といった対象テレビ番組における特徴的なキーワードを集めた辞書である. 辞書の作成方法としては, 予めすべての番組に対して考え得るすべてのキーワードを列挙して辞書を構築することも可能である. しかし, 莫大な辞書となるため現実的ではなく, 多くのメモリを消費してしまう. このため, 番組のTSに含まれているSI(Service Information)情報やインターネット等で提供されている番組情報を参照して本辞書を自動生成する.

表3-3 動画固有辞書の例

#	検出キーワード	提示キーワード
1	AKB47	出演者 (AKB47)
2	SMEP	出演者 (SMEP)
3	会いたかったな	曲名 (会いたかったな)
4	夜空ノコッチ	曲名 (夜空ノコッチ)
...

特に、インデクシング実施時に、当該インデクシング対象のテレビ番組に登場する出演者名や曲名及びそれらの通称を取得して、当該番組に特化した辞書を生成する。

先と同様に本辞書の音楽番組での利用方法を例示する。表3-3に音楽番組用の動画固有辞書を示す。検出キーワードとして出演者名と曲名を設定した例である。字幕情報内に検出キーワードの文字列が現れた時刻に、提示キーワードを用いてインデクシングする。これにより、出演者名や曲名が喋られた位置にインデックスを付けることができる。このため、ユーザは出演者名や曲名といった番組に対して固有なシーンの検索を実施可能となる。

本動画固有辞書は、定型句辞書による一般的なキーワードを意味的に詳細化したレベルのキーワードを持つ。このため、インデックス自体は少数であり、ユーザは好きなシーンを検索しやすくなるという長所がある。しかし、それらのインデックス位置はトピックの開始位置とは異なり、前後にズレる（検索位置ズレが生じる）場合が多い。例えば「出演者」や「曲名」等は、これらの言葉が単に喋られただけの場合も多く、動画固有辞書によるインデックス位置は正確ではない場合が多いという短所がある。

3.3.4 多重インデクシング法

本方式では，動画固有辞書によるインデックス位置のズレを補正する．これを実現するため，定型句辞書によるインデクシングと動画固有辞書によるインデクシングを行い，これらの結果を組合わせて最終的なインデックスを生成する(図3-5)．

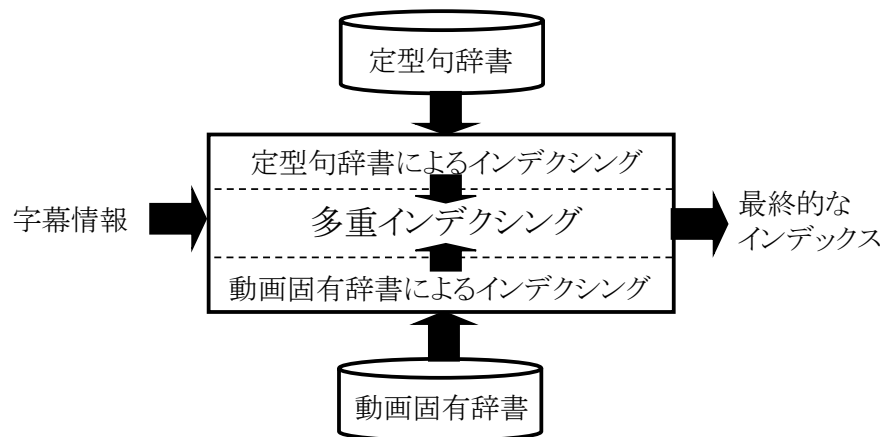


図3-5 多重インデクシング方式の構成

具体的には，定型句辞書及び動画固有辞書の検出キーワードが字幕情報内に出現した場合に，その字幕情報の表示時刻と，各辞書における提示キーワードを用いてインデクシングを行う．前述のとおり，定型句辞書によるインデックス位置は正確であるという長所がある．これを考慮し，動画固有辞書によるインデクシング結果を定型句辞書によるインデクシング結果に合わせ込む．

図3-6に音楽番組での多重インデクシング例を示す．番組開始から30秒の時点で「曲です．」という字幕が表示された場合，定型句辞書により時刻30秒の位置に「曲」という提示キーワードでインデクシングする．同様に，番組開始から10秒の時点で「AKB47」という字幕が表示された場合，動画固有辞

書により時刻10秒の位置に「出演者 (AKB47)」という提示キーワードでインデクシングする．そして，提示キーワード「出演者 (AKB47)」のインデックス時刻(10秒)をその直後の「曲」のインデックス時刻(30秒)に補正する．これにより，AKB47が紹介されただけの時刻を，実際にAKB47が曲を歌うシーンの時刻に合わせることができる．

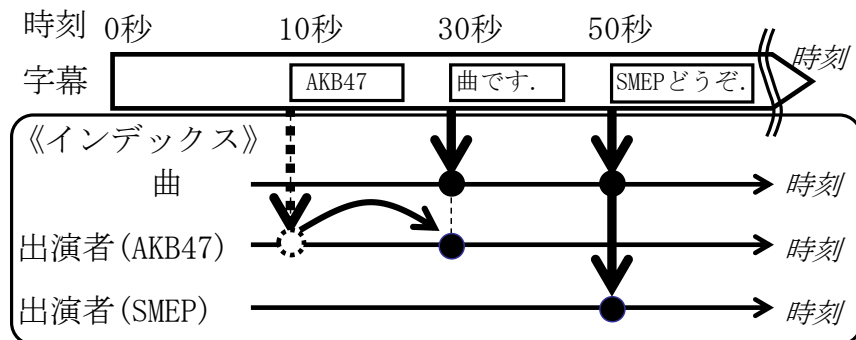


図3-6 多重インデクシングの例

上述の通り，本方式ではユーザにより指定されたキーワードのシーンが単に話題に挙がる時刻と，トピックの開始位置の時刻とのズレ(検索位置ズレ)を補正し，シーン検索精度を向上できる．また，シーンに複数の意味を持たせることができるため，ユーザは単に曲をザッピングして視聴するだけでなく，出演者名や曲名を指定して，好みの曲を視聴することができる．

3.4 テレビ向けシーン検索機能の実現

提案方式に基づいて，テレビ録画番組のシーン検索機能を実現した．本節では，特に，テレビ向けに実現した字幕符号化方式，辞書構築方法，字幕処理ルール，及びシーン検索機能の実現エンジンについて述べる．

3.4.1 字幕符号化方式

提案方式ではテキストである字幕情報を用いるため、映像や音声进行处理する場合に比べ、処理負荷を低減し、必要計算リソースを軽減することができる。これに加え、以下の字幕符号化方式を実現した。これにより、さらなる低処理負荷化を実現し、必要計算リソースの低減することが可能となる。以下では本字幕符号化方式について詳述する。

まず、動画固有辞書の登録キーワード数を最大64個とし、一つのキーワードあたりの文字数を14文字に制限する。通常、全番組に共通な辞書として、人名と曲名を登録する場合、少なくとも8万語程度のキーワードを持つ必要がある。この場合、1キーワード当たり14文字(符号値4バイトを合わせて32Byte)とすると、2.56MByte (8万語×32Byte)のメモリが必要となる。一方、上述の通り、登録キーワード数を最大64個とすることで、辞書に必要なメモリを2KByte程度 (64×32Byte)に削減できる。

次に、字幕パケットをすべて保持して処理するのではなく、インデクシングに必要な字幕パケットのみを定型句辞書及び動画固有辞書を用いて一旦符号化する。具体的には、2つの辞書における各検出キーワードを、辞書内での位置を表す番号に符号化し、表3-4に示す字幕特徴データベースを生成する。より詳細には、字幕パケット中に定型句辞書の検出キーワードが含まれている場合は、そのパケットの時刻情報と共に、対応する検出キーワードの番号を字幕特徴データベースの定型句辞書符号エリアに書き込む。同様に、字幕パケット中に動画固有辞書の検出キーワードが含まれている場合は、対応する検出キーワードの番号を固有辞書符号エリアに書き込む。字幕パケット中にこれらの検出キーワードが含まれていない場合は「0」を書き込む。定型句辞書符号エリア及び固有辞書符号エリアの両方が「0」の場合は、そのレコード自体を書き込まない。

表3-4 字幕特徴データベース

時刻(秒)	定型句辞書符号	固有辞書符号
10	0	1
30	1	0
50	5	2
...

字幕特徴データベースを用いた多重インデクシング法は以下の手順となる。まず、字幕特徴データベースを参照し、各時刻における符号値と一致する定型句辞書及び動画固有辞書の提示キーワードを用いて、それぞれインデクシングする。表3-4の字幕特徴データベースの例では、番組開始後「10秒」の時刻に固有辞書符号の符号値「1」が格納されている。この符号値「1」に対応する動画固有辞書の提示キーワード「出演者 (AKB47)」で時刻「10秒」のシーンをインデクシングする(表3-5)。同様に、番組開始後「30秒」の時刻に定型句辞書符号の符号値「1」が格納されている。この符号値「1」に対応する定型句辞書の提示キーワード「曲」で時刻「30秒」のシーンをインデクシングする(表3-6)。次に、動画固有辞書符号に基づく提示キーワードの時刻を、その前後の定型句辞書符号に基づく提示キーワードの時刻に補正する(図3-7)。

表3-5 動画固有辞書によるインデクシング

提示キーワード	時刻
出演者 (AKB47)	10
出演者 (SMEP)	50
・	・
・	・
・	・

表3-6 定型句辞書によるインデクシング

提示キーワード	時刻
曲	30
曲	50
・	・
・	・
・	・

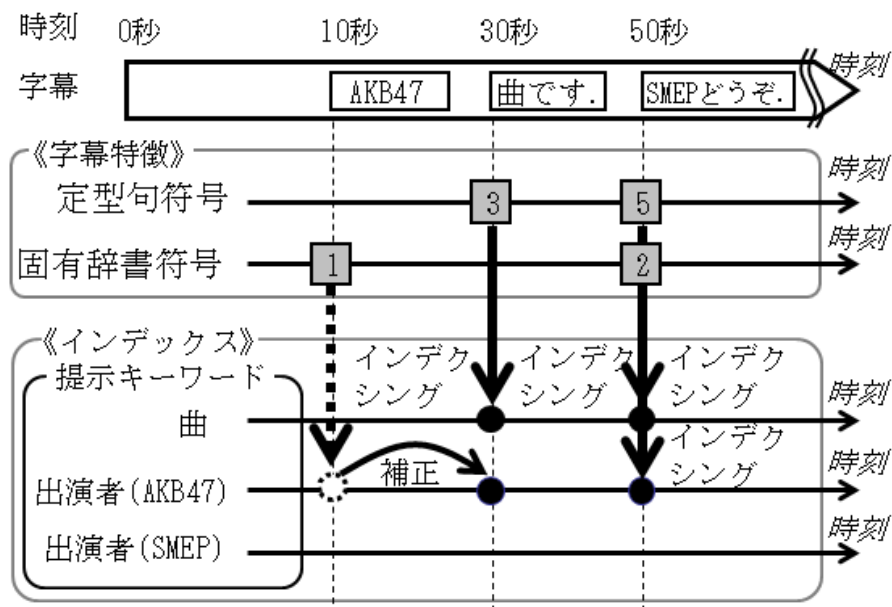


図3-7 インデクシングと時刻の補正

これにより、3.3節で述べた通り、動画固有辞書におけるキーワードのシーン時刻のズレを補正できる。

このようにして、特定の時刻のシーンを、提示キーワードで多重インデクシングした結果のインデックスデータを表3-7に示す。なお、表3-7において、位置数は1つのキーワードが出現した時刻数を示し、時刻はそのキーワードの出現時刻を示している。また、動画固有辞書におけるキーワード「出演者

(AKB47)」の時刻が「10秒」から「30秒」に補正された結果を示している。

本字幕特徴データベースのレコード長としては、時刻情報格納エリアを8Byte、定型句辞書符号格納エリアを1Byte、固有辞書符号格納エリアを1Byteとすることができる。実際には、今後の拡張を考慮して、各符号格納エリアを2Byteとし、4Byteの予約エリアを含め、1レコードのサイズを16Byteとした。

これらにより、辞書におけるデータサイズと、処理及び保持するデータ量を削減し、処理負荷と使用メモリ量を軽減することができる。

表3-7 インデックスデータ

提示キーワード	位置数	時刻
曲	2	30, 50
出演者(AKB47)	1	30
出演者(SMEP)	1	50
・	・	・
・	・	・
・	・	・

3.4.2 字幕処理ルール

提案方式では多重インデクシング法によりシーン検出の高精度化を実現している。しかしながら、シーンの検出精度を更に向上させるため、精度劣化の要因を検討し、これを改善する字幕処理ルールを構築した。

3.4.2.1 精度劣化の要因

以下にシーン検出精度の劣化要因を列挙する。

(1) 字幕処理単位

字幕解析において、通常では、放送局から送出される字幕を1パケット(1度に画面上に表示される表示単位)ごとに解析する。しかしながら、特定の単語がパケットをまたがって送出されるケースもある。例えば、野球番組などのリアルタイム性のある番組では、放送局における字幕生成者が番組内でしゃべられている言葉をリアルタイムに字幕として入力するため、この入力に間に合わず、単語が分割されてパケット化されるケースもある。例えば、「ホームラン」という単語が「ホーム」と「ラン」に分割されてパケット化されるケースも多く見られる。これらの単語が別のパケットにある場合には、それを判断できないため、シーン検出精度劣化の原因となる。

(2) 字幕特性

提案方式では、番組ジャンル毎に定型句辞書を使用して、字幕から一般的なキーワードが見つかった場合に、それらが見つかった位置を対象シーンとして検出する。しかし、実際には、番組毎に特定の言い回しや記号が使われていることも多い。例えば、サッカー番組では、実際にはアナウンサーは、「○○のシュートです。」とアナウンスしているが、ある番組では、そのまま字幕化され、別の番組では「○○のシュートでした。」と字幕化される。そのため、「シュートです」というキーワードを元に特定シーンを検出すると、ある番組ではこのシーンを全く検出できない事態が起こる。他方、「シュートでした」をキーワードとすると、「シュートです」を字幕化する番組にとっては、「シュートでした」は実際には過去に起こった出来事を字幕化しているため、無意味なシーンを検出することとなる。これは、ジャンルごとに決められたキーワードや記号を固定化しては、適合する番組と適合しない番組が出てきてしまい、番組毎に精度が良い番組と、そうでない番組が存在することを意味する。そのため、実際には番組ごとによく使われる言い回しやキーワードを認識して、これを用いて特定シーンを検出する必要がある。

(3) 対象シーン範囲

サッカーや野球の番組では、番組が始まってから実際に試合が始まるまでの間に試合前の会場の様子等が放送されている場合や、試合後にもハイライトやインタビュー等が放送されている場合もある。したがって、番組全体を対象として特定シーンを検出する場合、これらの無効シーンも検出することとなり、シーン検出精度が劣化する。そのため、これらの番組では、実際の試合のシーンのみを有効シーンとして検出し、この有効シーンから特定シーンを見つけることが必要となる。

(4) 単一キーワードでの判断

定型句辞書によるインデクシングでは、定型句辞書を使用して、「続いては」等の文脈の切り替わりを意味する単語を検出することにより、文脈の切り替わりを判断する。しかし、これらの単語は必ずしも（たとえアナウンサーが発していても）字幕化されないこともあり、その場合には、文脈の変化を検出できない。そのため、これらの単語だけに頼るのではなく、文脈の切り替わり点を判断することが重要となる。

(5) キーワードの不特定性

文脈の変化を検出して、シーンを特定しても、そのシーンに対するキーワードを特定できない場合には、当該シーンのキーワードをユーザに提示できない。一方で、1つのシーンに対してキーワードが複数含まれる場合や、逆にキーワードが全く含まれていない場合もある。このため、キーワードを唯一に特定できない場合もあり、これが特定シーンの検出精度劣化の原因となる。そのため、文脈の切り替わりを判断した後に、そのシーンの代表キーワードを的確に抽出あるいは生成することが必要となる。

3.4.2.2 特定シーンの抽出方式

本節では、前節で検討した精度劣化の要因を加味し、特定シーン抽出精度を向上させる方法を提案する。

(1) マルチ字幕パッケージ処理方式

本方式では、複数の字幕パッケージを統合し、複数文を用いて特定シーンを抽出する。具体的には、以下の字幕データの文字列の区切りルールに従って、1パッケージ単位で受け取る字幕データを複数でマージし、このマージされた字幕データを基にインデクシングを行う。

(a) 異なる色コードの字幕パッケージが出現した場合に区切る

話者の切り替わり時には字幕色が変更されることがあるので、字幕の色コードが変わる前の字幕パッケージの末尾を区切りとする。

(b) パッケージ時刻が一定時間以上離れた場合に区切る

パッケージの時刻が離れている場合、ここで話題が切り替わっている可能性が高い。このため、所定の時間（10秒）経過前の字幕パッケージの末尾を区切りとする。

(c) パッケージ数が予め定めた上限値に到達した場合に区切る

実際には区切らないことが望ましいが、使用メモリ量を考慮して、予め定めた数のパッケージ数（最大3パッケージ）で区切る。この場合、“。”、“！”，“？”等の「区切り文字」の直後を区切りとする。

(d) マージ字幕の文字数が予め定めた上限値を超えた場合に区切る

本ケースでも実際には区切らないことが望ましいが、使用メモリ量を考慮

して、予め定めた文字数（最大200文字）で区切る．この場合，“。”，“！”，“？”等の「区切り文字」の直後を区切りとする．

(e) 問かけ文を含む場合に区切る

このケースでは話者が切り替わることが多いため、問かけ文の末尾を区切りとする．具体的には“？”の直後を区切りとする．

(f) 話者表現を含む場合に区切る

このケースでも話者が切り替わることが多いため、話者表現を含む文の末尾を区切りとする．具体的には，“話者≫”，“話者>>”，“（話者）”等の話者表現が現れた場合に、それらの文字列の直前を区切りとする．

(2) 番組適応字幕解析方式

本方式では、番組の字幕特性を学習し、番組固有の記号解析及び字幕解析を実施する．具体的には、以下のルールに従って、字幕データの選別処理を行う．

(a) 特定シーンの抽出数と出現数閾値との比較演算により、特定シーンの有効の可否を判定する

本ルールは野球，サッカー，ニュースに適用する．例えば、字幕内に「！」の出現数が多い場合はおそらくノイズであるが、番組によっては「先制点！」などの特徴的なシーンにおいてのみ使用される場合があり、特定シーンとして利用できる．

(b) 特定シーンの抽出数と番組情報から抽出したキーワード出現数との比較演算により、特定シーンの有効の可否を判定する

本ルールは音楽，バラエティ番組に適用する．具体的には、曲区切りの判

定キーワードを複数設定し、抽出数が出演者名数以上のものは排除する。

(3) 文脈変化検出方式

本方式では、文脈（話題）の変化点を検出することでトピックの先頭を抽出する。具体的には、以下のルールに従い、字幕の文脈が変化する位置を判定する。

- (a) 字幕の処理単位での使用単語を比較し、近接の字幕間で使用単語に重複がない場合に話題が変化した（文脈が変化した）と判定する

本ルールはニュース番組に適用し、字幕に含まれる任意の単語が、処理単位での直前の字幕には含まれていない場合に話題が変化したと判別する。ここで、単語とは「漢字・カタカナ」の文字列を切り出したものとする。

- (b) 字幕の色や表示位置が変化した場合に、前後の字幕との時間間隔が長ければ、話題が変わったと判断する

本ルールはバラエティ番組に適用し、色や表示位置が変化した場合に、その前後で時間が空いていた場合には、話題が切り替わったと判断する。

- (c) バラエティ番組において、短時間内に異なる複数の人名キーワード（動画固有辞書に人名として登録されているキーワード）を検出する場合に以下のルールで話題の変化を判別する

ルール1：字幕内で検出した人名数に応じてインデクシングを行うか否かを決定する。特に、人名数が少ない場合は、インデクシングを行う。これにより、例えば、複数の出演者が同じタイミングで演じている場合に対応できる。一方、人名数が多い場合は、インデクシングを行わない。これにより、例えば、本人不在でコメント内で人名を言及されている場合に対応できる。

ルール2：字幕内で検出した人名キーワードが「司会者」であった場合，文脈の変化点と判定する．これは，例えば，字幕内で明確にトピックが示されていない場合に有効となる．なお，人名キーワードが司会者であるか否かは例えば番組情報から判別することが可能である．

ルール3：字幕内で検出した人名キーワードに「出演者名」以外の人名が入っている場合は文脈の変化点と判定しない．これは，例えば，出演者ではなくゲストがコメントしている場合に対応する．

(4) 代表キーワード抽出方式

本方式では，シーンに関連するキーワードを抽出し，代表キーワードとして提示する．具体的には，定型句辞書と動画固有辞書のそれぞれの辞書によるインデクシング時及び多重インデクシング時に以下の処理を行う．

<各辞書によるインデクシング時>

各辞書によるインデクシング時には，以下のルールに従い，字幕文字列内の特徴的なキーワードを検出し，インデクシングを実施する．

(a) 番組情報から人名を抽出し動画固有辞書へ登録する場合に，出演者名と，それ以外の人名（司会，ゲスト，進行など）とを区別して登録する

本処理はバラエティ番組に適用する．これにより，例えば，番組に出演している「芸人」と，それ以外の出演者を（司会，ゲスト，進行など）とを区別できる．なお，出演者が司会であるか，ゲストや進行であるかは番組情報に記載されているため，この情報を用いて判断することができる．

(b) 字幕パケット内に複数のキーワードが含まれる場合，全てのキーワードを用いてインデクシングを行う

本処理はバラエティ番組に適用する．これにより，字幕パケット内の複数キーワードを列挙可能となる．

- (c) 音楽番組において，楽曲名を示すキーワードとみなす文字列を字幕文字列から切り出し，楽曲名を示すキーワードとする．具体的には，鉤括弧（“「”と“””）で囲まれたキーワードを曲名キーワード候補として出力する（ただし番組名と一致するキーワードは除く）．

- (d) 音楽及びバラエティ番組において，出演者名を示すキーワードとみなす文字列を以下のルールで字幕文字列から切り出し，出演者のシーンをインデクシングする．

ルール1：「～です。○○さんの歌声が～」の「○○」等，文区切り（句読点，感嘆符，空白）と敬称（“さん”，“君”，“ちゃん”）に囲まれたキーワードを出演者キーワードとする．

ルール2：「ここからが勝負だ。□□○○。」の「□□○○」等，動画固有辞書に含まれる出演者名キーワードの直後に“。”が出現する場合，出演者キーワードに加えてトピックキーワードとしても抽出してインデクシングを行う．

<多重インデクシング時>

音楽番組及びバラエティ番組では，定型句辞書及び動画固有辞書における抽出キーワードと連動して，以下のルールに従ってインデクシングを行う．

- (a) 音楽番組におけるインデクシングルール

ルール1：動画固有辞書に含まれる出演者名の数が1名の場合に，曲区切りに対応する出演者名キーワードをすべて同一の出演者名とする．これはある番組が特定のアーティストに対する特番で

ある場合に有効となる。

ルール 2 : 曲区切り間で複数の出演者名キーワードが検出されている場合、最も検出数の多いものを出演者名キーワードとする。

ルール 3 : 曲区切り間で複数の曲名キーワードが検出されている場合、最も検出数の多いものを曲名キーワードとする。

ルール 4 : 曲区切りの位置に到達前に、同名キーワードが重複して検出された場合、最初のキーワードのみをインデクシングに採用する。

ルール 5 : 曲区切りの位置に到達前に、複数の異なるキーワードが検出された場合、最初に検出した（曲区切り位置から最も遠い位置の）キーワードのみをインデクシングし、残りのキーワードは無視する。

ルール 6 : 曲名や出演者名キーワードが連続して検出された場合で、同名のキーワードが連続する場合には、重複キーワードとして最初のキーワードのみを採用する。一方、異名のキーワードが連続する場合には、連続して楽曲を流すケースに対応するため、それぞれが最も近い位置の曲区切り（ただし他のキーワードと位置あわせが重複していない）の位置に合わせる。

ルール 7 : 動画固有辞書の曲名キーワードの数と曲区切りの数が等しく、曲名キーワードの出現順が番組内の曲順序と一致する場合に、曲名キーワードと関連付けられていない（位置合わせされていない）曲区切りが存在する時は、曲名キーワードの出現順序に従ってこれらを関連付けることで補完する。

ルール 8 : 動画固有辞書の出演者名キーワードの数と曲名キーワードの数が等しく、曲名キーワードの出現順が曲区切りに関連付けられている曲名キーワードの順と一致する場合に、出演者名キーワードと関連付けられていない（位置合わせされていない）曲

区切りが存在する場合は，出演者名キーワードの出現順序に従ってこれらを関連付けることによって補完する．

(b) バラエティ番組におけるインデクシングルール

ルール1：あるトピックから次のトピックまでに，出演者名キーワードが連続して検出された場合の動作は以下の通りとする．

- (i) 同名キーワードが連続：次のトピックの位置に到達前に，同名キーワードが重複して検出された場合，最初のキーワードのみをインデクシングに採用する．
- (ii) 異名キーワードが連続：次のトピックの位置に到達前に，複数の異なるキーワードが検出された場合，全てのキーワードを有効にする．これにより，複数人が同時に出演している場合に対応できる．

ルール2：動画固有辞書の出演者名キーワードの数とトピックの数が等しく，出演者名キーワードと関連付けられていない（位置合わせされていない）トピックが存在する場合は，出演者名キーワードの出現順序に従ってこれらを関連付けることによって補完する．

3.4.3 辞書構築方法

提案方式では動画固有辞書を番組情報に基づいて生成する．しかし，普及型の録画機能付きテレビは必ずしもインターネット等に接続されていない可能性があること及びテレビは誰にでも使える機能を提供する必要があることを考慮して，本実現においては，動画固有辞書を基本的にSI(Service Information)情報に含まれる番組情報に基づいて自動生成する構成とした．SI情報はテレビ番組のTSに含まれている．このSI情報の内容記述セクションに

は、番組の出演者名や司会者名、番組内で歌われる曲名等の、番組に固有の情報がテキストとして含まれている(図3-8)。本実現では、これらの情報を検出キーワードとして取得し、動画固有辞書を生成する。そして、テレビがインターネットに接続されている場合には、サーバ上に構築した通称辞書から出演者名や曲名の通称を取得し、動画固有辞書の検出キーワードとして登録できるように構成した。

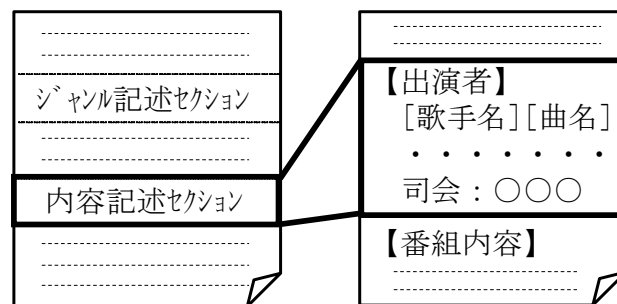


図3-8 SI情報とキーワード

3.4.4 シーン検索機能実現エンジン

提案方式に基づいて、シーン検索機能を実現した。本実現エンジンの構成を図3-9に示す。

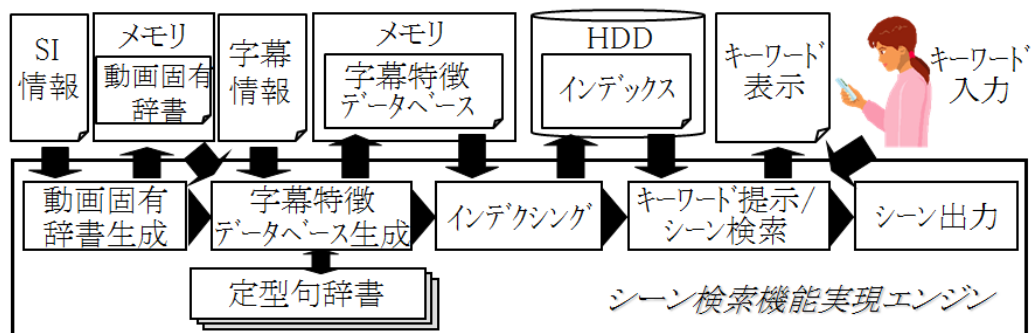


図3-9 シーン検索機能実現エンジンの構成

本エンジンは以下の動作を行う。

- ① テレビ番組の録画開始時に番組情報を取得して動画固有辞書を生成する。
- ② 録画中に字幕情報を取得し、定型句辞書と動画固有辞書に基づいて、字幕特徴データベースを生成する。
- ③ 録画終了後に定型句辞書及び動画固有辞書によるインデクシングと多重インデクシングを行う。
- ④ ユーザがテレビ番組を再生する際やキーワードを入力する際に、インデックスに含まれる提示キーワードを図3-10に示すユーザインタフェースの通り表示する。
- ⑤ ユーザからキーワードが選択された場合、そのキーワードのシーンの時刻をチャプタ点として表示する。

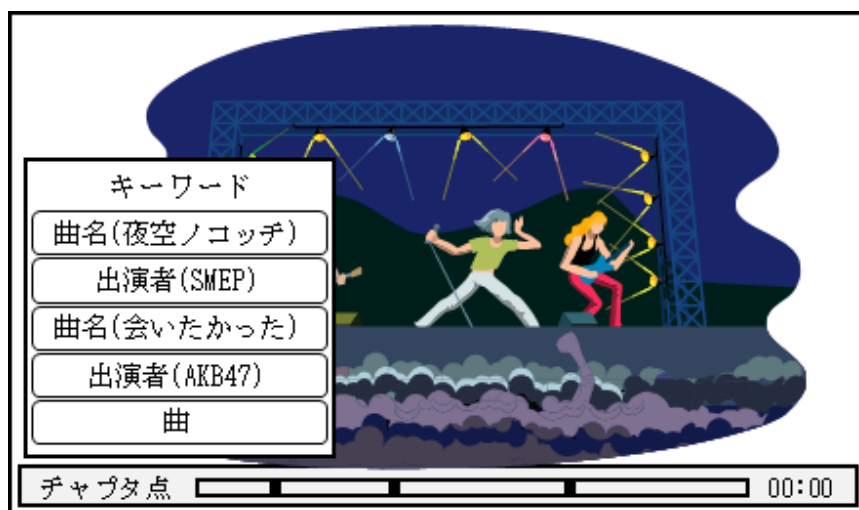


図3-10 シーン検索機能のユーザインタフェース

これにより、ユーザは提示されたキーワードをリモコンで選ぶだけで、録画テレビ番組の観たいシーンを簡単に視聴可能となる。なお、ユーザに提示

されるキーワードは、先の定型句辞書及び動画固有辞書で示した提示キーワードである。本エンジンでは、テレビ番組のジャンル毎に表3-8のチャプタ位置に提示キーワードを設定できる辞書を構成するようにした。これにより、ユーザは提示されたキーワードを見るだけで、録画番組の視聴前に番組に含まれる曲やホームランシーンの有無等を把握可能とした。

表3-8 チャプタ位置及び提示キーワード

#	ジャンル	提示キーワード	チャプタ位置
1	音楽	曲	曲の始まり
		<出演者名>	<出演者名>の曲の始まり
		<曲名>	<曲名>の曲の始まり
2	ニュース	トピック	トピックの始まり
		お天気	天気の話の始まり
		スポーツ	スポーツの話の始まり
3	バラエティ	トピック	トピックの始まり
		<出演者名>	<出演者名>が出るトピックの始まり
4	野球	みどころ	ホームラン、得点、ファインプレーのシーン
		ホームラン	ホームランシーン
		得点	得点シーン
		ファインプレー	ファインプレーシーン
5	サッカー	みどころ	キックオフ、ゴール、シュート、前後半終了シーン
		ゴール	ゴールシーン
		シュート	シュートシーン
		キックオフ	キックオフのシーン
		前半／後半終了	前半／後半終了のシーン

3.5 提案技術の評価と関連技術との比較

実際のテレビ番組を使用して、本提案技術におけるシーン検索精度と処理負荷を評価した。処理負荷としては処理に必要な計算リソース（CPU負荷と使用メモリ量）を評価した。また、これらについて関連技術と比較した。以下にこれらの評価結果及び比較結果を示し、考察する。

3.5.1 提案技術の評価

本提案技術によるシーン検索機能を1.40GHzで動作するIntel社製のCeleronプロセッサM360及び256MByteのメモリを搭載したWindowsPC(以下,実験PCと呼ぶ)上で動作させ,シーン検索精度を評価した.

実験では,字幕情報が付与されている番組を対象とした.実験番組数は表3-8に記載のジャンル毎に42番組,合計210番組である.これらの番組に対して,本実現したシーン検索機能によりチャプタを表示し,表示チャプタ数(Pn)と表示チャプタ位置(Pp)を得た.

評価では,検索結果の正しさを示す適合率と,検索の網羅性を示す再現率を用いる.適合率と再現率は次の通り表される.

$$Pr = n(D \cap R) / n(R) \quad (3.1)$$

$$Re = n(D \cap R) / n(D) \quad (3.2)$$

ここで, Pr と Re は,それぞれ適合率と再現率を示す.また, D と R は,それぞれユーザが望むチャプタの集合と検索されたチャプタの集合を示している.特に,本評価では,実験で用いた210番組のジャンルに合わせて,表3-8で定義した提示キーワードシーンのトピック開始位置を手で列挙し,理想チャプタ数(In)と理想チャプタ位置(Ip)を得た.そして,各提示キーワードシーンのトピック開始位置に対する理想チャプタ位置(Ip)を中心として,±5秒以内にそのキーワードシーンの表示チャプタ位置(Pp)が入っていた場合に,この表示チャプタ位置を正解とし,正解チャプタ数(An)を得た.

シーン検索精度として,番組ジャンル毎に正解チャプタ数(An)と表示チャプタ数(Pn)及び理想チャプタ数(In)を集計した上で,理想チャプタ数(In)で正規化して,(3.3)式及び(3.4)により,それぞれ適合率(Pr)と再現率(Re)を算出した.結果,シーン検索適合率及び再現率は,それぞれ80%及び83%であ

った．このことより，(3.5)式で示される適合率と再現率を総合的に示すF値を81.5%とすることができた．

$$Pr = An / Pn \quad (3.3)$$

$$Re = An / In \quad (3.4)$$

$$F\text{値} = 2 * Pr * Re / (Pr + Re) \quad (3.5)$$

CPU負荷の評価にも先の210番組を用いた．具体的には，上述した実験PCにおいて，各番組の録画中に行うインデクシング処理に必要な処理時間を測定した．結果，各番組において，番組が1時間（ $=3.6 \times 10^6$ ミリ秒）とした場合の処理時間は最大でも4719ミリ秒であった．実験PCのCPUは1.40GHz（ $=1400\text{MHz}$ ）で動作しているため，CPU負荷は次式のとおり1.835MHzとなる．

$$1400\text{MHz} \times 4719\text{ミリ秒} \div (3.6 \times 10^6\text{ミリ秒}) \div 1.835\text{MHz} \quad (3.6)$$

使用メモリ量に関しては，インデクシング処理において理論上最大で使用するメモリ量を計算した．結果，必要となる最大使用メモリ量は，辞書領域として4Kbyte（動画固有辞書2KBと定型句辞書2KByte），字幕特徴データベースの領域として128Kbyte，作業領域として252Kbyteの合計384Kbyte（ $=0.384\text{MByte}$ ）であった．このうち，字幕特徴データベースは番組の長さに依存する．前述のとおり，字幕特徴データベースのレコード長は16Byteであるため，字幕特徴データベースのメモリ領域を128Kbyteとすることで，最大レコード数を8192とすることができる．これは，仮に12秒に1度の間隔で動画固有辞書あるいは定型句辞書のキーワードが出現した場合でも，98304秒すなわち27.3時間の番組に対応できることを意味する．実際には1つの番組内で12秒に1度の間隔でコンスタントにトピックの変化点を示す定型句辞書のキーワードや曲名等の動画固有辞書のキーワードが出現することは考えにくい．このため，計算したメモリ量で近年の最長番組である27時間テレビにも対応できるといえる．

3.5.2 関連技術との比較

比較として、関連技術であるテロップ文字認識による方式[23]、顔認識による方式[29]、音声認識による方式[34]、字幕解析による方式[39]、テロップ文字認識と音声認識及び字幕解析を利用した複合方式[49]を評価した。精度の評価では、提案技術の評価で用いた210番組の中から各ジャンルに対して5番組ずつ抽出した合計25番組を用いて、すべての方式を同じ番組で評価した。特に、本評価では各方式における技術で公開されている認識率に基づいて、提案機能と同一の機能を実現するものとしてシミュレーションを行った。シミュレーションでは、まず、評価に用いる25番組の各番組において、テロップや顔が出現する時間位置及び音声や字幕内でキーワードが出現する時間位置をそれぞれ人手で列挙した。次に、これらの時間位置が先の提案技術の評価と同様の基準でチャプタ位置として正しいか否かを人手で判断することで、それぞれの方式における再現率と適合率を算出した。さらに、これらの再現率と適合率にそれぞれの方式における最大の認識率を乗じることで、最終的な再現率と適合率を算出し、(3.5)式によりF値を算出した。結果、F値は、テロップ文字認識による方式で71.0%(適合率87%, 再現率60%)、顔認識による方式で37.3%(適合率43%, 再現率33%)、音声認識による方式で71.9%(適合率64%, 再現率82%)、字幕応用による方式で65.9%(適合率56%, 再現率80%)、テロップ認識と音声認識及び字幕応用を利用した複合方式で89.4%(適合率93%, 再現率86%)となった。

CPU負荷については、テロップ文字認識による方式[17-23]、顔認識による方式[24-29]、音声認識による方式[30-34]、字幕解析による方式[35-41]のそれぞれで公開されているインデクシング処理に必要な最小のCPU負荷を用いた。また、複合方式についてはテロップ文字認識、音声認識、字幕解析のそれぞれでインデクシング処理に必要となるCPU負荷を積み上げることで計算

した．結果，最小のCPU負荷は，テロップ文字認識による方式で200MHz，顔認識による方式で240MHz，音声認識による方式で60MHz，字幕応用による方式で8MHz，テロップ認識と音声認識及び字幕応用を利用した複合方式で268MHzであった．

使用メモリ量については，テロップ文字認識による方式[17-23]，顔認識による方式[24-29]，音声認識による方式[30-34]，字幕解析による方式[35-41]のそれぞれで評価した．特に，関連技術の節で述べたとおり，各関連技術でのインデクシング処理に必要なデータサイズと辞書サイズに基づいて，作業領域と辞書領域として最低限必要となるメモリ量を算出した．複合方式についてはテロップ文字認識，音声認識，字幕解析のそれぞれでインデクシング処理に最低限必要となるメモリ量を積み上げることで算出した．実際には，提案技術の評価で考慮したとおり，シーンの特徴データあるいはインデックスデータを保持する領域が必要となる．しかし，この領域は，作業領域と辞書領域に必要なメモリ量に比べて非常に小さい．また，本関連技術の評価では少なくとも必要となるメモリ量が提案技術に比べて十分に大きいことが確認できれば良い．これらのことから，本関連技術の評価ではシーンの特徴データやインデックスデータの保持に必要なメモリ量を無視することができる．これらの条件のもとで，関連技術による方式のそれぞれにおいてインデクシング処理に少なくとも必要となるメモリ量を算出した．結果，少なくとも必要となる最小のメモリ量は，テロップ文字認識による方式で6MByte，顔認識による方式で6MByte，音声認識による方式で3MByte，字幕応用による方式で2MByte，テロップ認識と音声認識及び字幕応用を利用した複合方式で7MByteであった．

提案技術についても，関連技術での評価と同じ25番組を用いて先の210番組による評価方法と同様の方法で精度を評価しなおした．結果，F値は81.0%(適合率80%，再現率82%)であった．一方で，使用計算リソースについては，25番組を用いても，先に示した210番組を用いて得られた使用計算リソース以上

になることはない。このため、提案技術で必要となる最大の使用計算リソースは、先に示したCPU負荷1.835MHz及び使用メモリ量0.384MByteであるといえる。したがって、これらの値に基づいて、(3.7)式により関連技術と提案技術に対する単位リソースあたりの精度(以下、リソース精度比と呼ぶ)を求めた。結果、リソース精度比は、提案方式で 1.15×10^2 、テロップ文字認識による方式で 5.91×10^{-2} 、顔認識による方式で 2.59×10^{-2} 、音声認識による方式で 3.99×10^{-1} 、字幕応用による方式で4.12、テロップ認識と音声認識及び字幕応用を利用した複合方式で 4.76×10^{-2} となった。

$$\text{リソース精度比} = F\text{値}(\%) / \text{CPU負荷}(\text{MHz}) / \text{メモリ量}(\text{MByte}) \quad (3.7)$$

これらによる各技術での精度(F値)とリソース精度比を図3-11に示す。

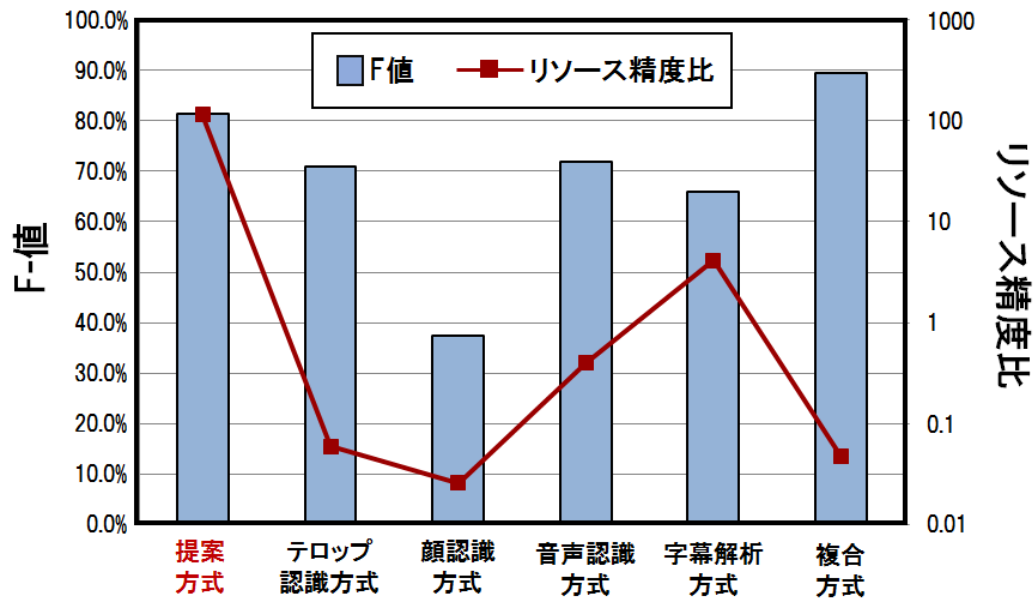


図3-11 関連方式との比較結果

3.5.3 評価結果及び比較結果に対する考察

本評価結果では、提案方式のシーン検索精度は複合方式の精度に次いで二番目となった。これは、字幕が表示されずにキーワードが生成できなかったシーンに対して、複合方式では音声認識やテロップ認識で生成できたためである。なお、顔認識による方式は人名のキーワードしか提示できないため、他の方式と比べて精度が低くなっている。また、音声認識の精度は100%ではないにも拘らず、音声認識による方式が字幕応用による方式に比べて精度が良い結果となっている。これは実際にはしゃべられている言葉が字幕では必ずしも表示されないことに起因している。さらに、テロップ文字認識による方式は音声認識による方式と比べて若干精度が劣っている。これは、認識精度が音声認識の方が良い一方でテロップの出現時刻がトピックの開始位置に現れることが多いことに起因している。提案方式がこれらの方式と比べて良い理由は多重インデクシングによってインデクシング位置が補正されていることと、認識誤りがないためであると言える。

以上の通り、提案方式のシーン検索精度は十分に高い一方で、処理負荷を軽減でき、必要となる計算リソースが最も少なく、リソース精度比は最良の値となっている。また、関連技術のいずれの方式も本研究で目的としたCPU負荷3MHz以下、メモリ量0.5MByte以下でのシーン検索機能の実現は不可能であるのに対して、提案方式ではCPU負荷1.835MHz、メモリ量0.384MByteと極めて少なくすることができ、本研究で目的とした3MHz以下、メモリ量0.5MByte以下でのシーン検索機能の実現が可能となる。

関連技術と提案技術における精度と必要計算リソースを第2章で述べた処理負荷と精度の位置づけの図にマッピングすると図3-12のとおりとなる。図3-12では、処理負荷として次式で表されるCPU負荷と使用メモリ量の積を用いた。これは(3.7)式の分母であり、必要計算リソースを示している。

$$\text{必要計算リソース} = \text{CPU負荷(MHz)} \times \text{メモリ量(MByte)} \quad (3.8)$$

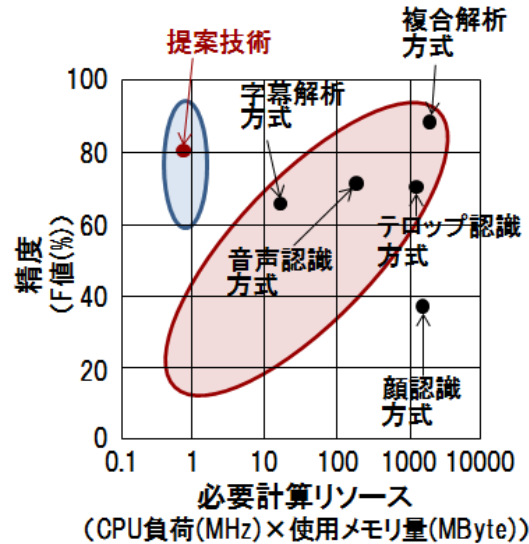


図3-12 処理負荷と精度の位置づけに対するマッピング

図3-12からわかるとおり，提案技術は他の関連技術では入っていない本研究のターゲットに入っており，必要計算リソースの低減と高精度化を実現できている．

以上により，普及型の録画機能付きテレビにおいて，CPUやメモリを変更及び追加搭載せず，ハードウェア構成を変えることなく，すなわちテレビの原価コストを上げることなく，シーン検索機能を実現することができた．特に，ユーザの指定したキーワードのトピックをズレなく検索可能とする（検索ズレのない）シーン検索機能を実現することができた．

なお，本提案方式では，以下によりさらなるシーン検出精度を向上することができると思われる．

(1) 字幕処理ルールの調整

本実現では，各番組ジャンルに対して20番組ずつ合計100番組を調査して，前節で示した字幕処理ルールを調整した．その結果，例えば野球のホームラ

ンや得点シーンの検出精度は3回の調整により図3-13に示すとおり改善していった。今後は、これらのルールにおけるパラメータの調整にタグチメソッド[50]を使用するなど、さらなる調整を行うことで、シーン検索精度が改善するものとする。

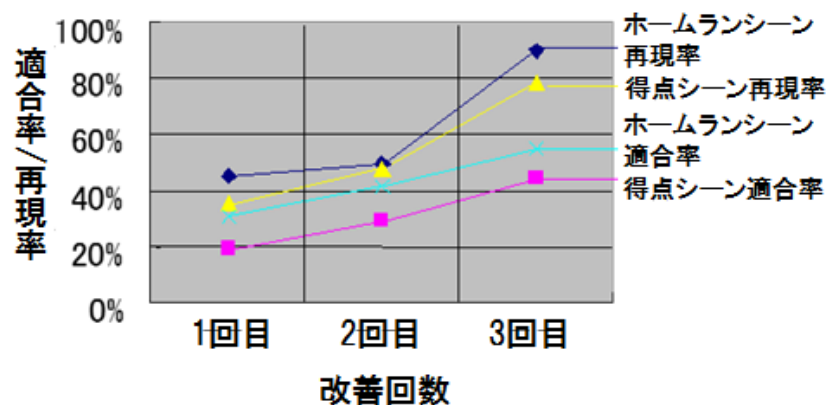


図3-13 野球のキーワードに対するシーン検出精度改善過程

(2) 生放送番組への対応

生放送番組では、実際に喋られた時刻から数秒遅れて字幕が表示される。本実現では、ニュース、野球、サッカーを生放送と考えて、これらの番組では字幕の表示時刻を一律で7秒早めて処理している。これに対して、生放送番組であることを番組ごとに正しく判断し、字幕の表示時刻を最適に調整することでシーン検索精度を向上することができる。よって、例えば番組情報から「生放送」等のキーワードを検出することで生放送番組であることを正確に判断し、音声と字幕のキーワードマッチング等により字幕の表示時刻を最適化する方法を検討していく必要がある。

(3) 多重インデクシングにおける時刻補正方法の改善

現状では、動画固有辞書におけるキーワードの出現時刻に対して±180秒以

内に定型句辞書のキーワードが出現していた場合に、動画固有辞書でのインデックス時刻を定型句辞書でのインデックス時刻に合わせ込んでいる。また、合わせ込む候補となる定型句辞書のインデックス時刻が複数存在する場合には、動画固有辞書におけるキーワードの字幕表示時刻に最も近い時刻に合わせ込んでいる。例えば図3-6において「AKB47」の字幕表示時刻の±180秒以内に「曲です。」の字幕が表示された場合に、「出演者(AKB47)」のインデックス時刻を「曲」のインデックス時刻に合わせ込んでいる。また、合わせ込む候補となる「曲」のインデックス時刻が複数存在する場合には、「AKB47」の字幕表示時刻に最も近い時刻に合わせ込んでいる。実際には、これらの条件をより細分/最適化した時刻補正方法によって、合わせ込む時刻を選定することで、シーン検索精度を向上できると考えられ、今後これらの時刻補正方法における条件の細分/最適化を検討していく。

3.6 新規性と有用性

本研究の新規性は字幕情報を用いた多重インデクシング方式とキーワードの符号化を行うところにあり、その有用性は字幕情報利用で発生し得る検索位置ズレを補正できる点と、インデクシングにおける処理負荷を軽減し必要計算リソースを低減できる点を両立できることにある。

従来、シーン検索機能を実現する技術としては、テロップ文字認識による技術、顔認識による技術、音声認識による技術、字幕情報解析による技術、及びこれらの組み合わせ解析による技術があり、テレビ番組におけるシーンのキーワードを生成し、シーンのインデクシングを行う技術があった。しかし、これらの技術では、いずれも処理負荷が高く、300MHzで動作するCPU及び256MByteのメモリしか搭載していない普及型の録画機能付きテレビにおいて、CPUやメモリを高性能なものに変更あるいは追加搭載することなく、シーン検索機能を実現することは困難であった。特に、テロップ文字認識による技術

では、テロップ（例えば人物名等）がトピックの変化点で表示されることが多いため、検索位置ズレの少ないシーン検索が可能となる。このため、シーン検索精度の面では有効であるが、処理負荷が非常に高く、多くの計算リソースを必要とする。字幕情報解析による技術は、他の技術と比べてシーン検索機能の実現における処理負荷を軽減し、必要となる計算リソースを低減できるが、それでも字幕情報の中から不定キーワードを抽出する必要があり、形態素解析等の自然言語処理で多くのCPU負荷を消費してしまっていた。また、シーン検索に有効なキーワード辞書に多くのメモリ量を必要としていた。このため、CPUやメモリを高性能なものに変更及び追加搭載することなく、普及型の録画機能付きテレビにシーン検索機能を付加することは困難であった。さらに、字幕情報は基本的には番組内でしゃべられている音声そのままテキスト化されているため、しゃべられるキーワード（例えば人物名等）は必ずしもトピックの最初だけに出てくるとは限らない。このため、トピックの開始位置に対する検索位置ズレが生じてしまい、シーン検索精度としては劣化してしまうという問題があった。

本研究では、テレビ番組のジャンル毎に共通のキーワードからなる定型句辞書と、番組に固有のキーワードからなる動画固有辞書の二種類の意味レベルの異なる辞書を使用した多重インデクシング法により、ユーザの指定したキーワードのトピックをズレなく検索可能とするシーン検索機能を実現可能とした。また、形態素解析等の自然言語処理を必要とせず、番組に対して必要最低限の辞書とキーワードの符号化により、シーンを効率的に表現することで、シーンインデクシングにおける処理負荷を軽減し、シーン検索機能の実現に必要な計算リソースを低減可能とした。

本研究成果により、他の技術では実現することが困難な図3-1の目標を達成した。これにより、CPUやメモリを変更及び追加搭載することなく、普及型の録画機能付きテレビにシーン検索機能を付加可能とした。すなわち、既存の普及型の録画機能付きテレビにおいて、ハードウェア構成を変えず、原価コ

ストを上げることなく，シーン検索機能を付加可能とした．

3.7 第3章のまとめ

第3章では，テレビ番組に付随する字幕情報を用いたシーンインデクシング方式を提案した．特に，テレビ番組のジャンル毎に共通なキーワードからなる定型句辞書と，テレビ番組に固有なキーワードからなる動画固有辞書を用いて，シーンの多重インデクシングを行うことで，シーン検索における適合率と再現率として，それぞれ80%及び83%(F値=81.5%)を達成した．また，キーワードを符号化し，シーンを効率的に表現することで，処理負荷を軽減し，CPU負荷を最大1.835MHz，メモリ使用量を最大0.384MByteとした．

本研究の新規性は多重インデクシング方式とキーワードの符号化を行うところにあり，その有効性は字幕情報利用で発生し得る検索位置ズレを補正できる点と処理負荷を軽減し必要計算リソースを低減できる点を両立できるところにある．本成果により，普及型の録画機能付きテレビにおいて，CPUやメモリを変更及び追加搭載せず，ハードウェア構成を変えることなく，すなわちテレビの原価コストを上げることなく，シーン検索機能を付加可能とした．

今後は，前節で示したパラメータ調整等を行うことにより，さらなるシーン検索精度の向上を図る．また，本研究により，テレビやハードディスクレコーダ等の端末側にキーワードを提示してシーンを検索するためのフレームワークを作成できたので，今後はメタデータ配信サービスあるいはメタデータ交換サービス等のサービスへの対応も検討していく．さらに，対応ジャンルの拡大とユーザ嗜好に基づいた，優先的なキーワード提示の実現および選択されたキーワードに基づくユーザ嗜好性の抽出とそれに基づくシーン検索/テレビ番組推薦の連携機能の実現により，さらなる使い勝手の向上を検討していく．

第4章

テレビ番組推薦技術

第3章では、テレビ番組を短時間で視聴するための技術として、ユーザが好みのシーンを選択して視聴可能とする「シーン検索機能」の実現技術について述べた。本章では、ユーザの好みに合ったテレビ番組を簡単に選択可能とし、見忘れないように推薦及び自動録画する「テレビ番組推薦機能」の実現技術を提案する。特に、推薦計算に必要となる処理負荷を軽減し、必要計算リソースを低減して、高精度なテレビ番組推薦を実現するために、ユーザの嗜好を的確に表現する必要最低限のテレビ操作と番組重要属性を定義する。具体的には、テレビ操作として、再生、予約、検索、録画を定義し、重視属性として、番組タイトル、番組キーワード、番組ジャンル、放送チャンネル、放送時間を定義する。そして、これらの重み付けにより、推薦対象のテレビ番組に対するユーザの嗜好度を示す推薦スコアの計算方式を提案する。また、推薦スコアと推薦アイテム数の関係を示す“Number-Score Curve (N-S Curve)”を用いた推薦アイテム数の決定方式を提案する。N-S Curveを用いて推薦アイテム数を決定することで、テレビ番組推薦における適合率と再現率のバランスを最適化し、限られた推薦番組数でユーザが好む番組を可能な限り過不足なく推薦可能とする。

4.1 はじめに

第1章で述べたとおり、近年、視聴可能なテレビ番組が大量化している。特に、テレビ番組に限らず、一般的な情報の大量化に伴い、ユーザは所望する情報を素早く取得することが困難になってきている。そのため、情報やテレビ番組をはじめとする「アイテム」をユーザの状況や好みに応じて推薦する「推薦システム」に対するニーズが高まっており、数々の推薦技術や推薦システムが提案されている。これらの推薦においては、ユーザに多くのアイテムを推薦した場合には、ユーザは多くの選択肢を得ることが可能となるが、好みではないアイテムも推薦される可能性があり、所望のアイテムを選択することが困難となる。一方で、少ないアイテムしか推薦しない場合には、ユーザはアイテムを選択することが容易となるが、選択肢が少なく、所望のアイテムが推薦されない可能性がある。すなわち、これらはトレードオフとなり、推薦においてはこれらのバランスが重要となる。特に、テレビはリモコンで遠くから操作されるため、番組選択のための視覚上の表示解像度や入力のためのユーザインターフェースが限られている。また、番組自動録画のためのハードディスク容量や記憶容量といったハードウェアリソースも限られている。このため、これらの機器におけるテレビ番組の推薦を信頼性があり、使い勝手の良いものとするためには、限られた推薦番組数で、ユーザが好む番組を可能な限り過不足なく推薦可能とする必要がある。これは推薦における適合率と再現率のバランスをとる必要があることを意味する。また、テレビ番組の推薦では、放送される番組や終了する番組が時々刻々と変化するため、ユーザからの要求に対して推薦番組をリアルタイムに計算する必要がある。可能な限り高速に推薦処理を実施する必要がある。具体的には、ユーザがあるリクエストをテレビに対して行った際に、ユーザが待つことができる時間は10秒程度であるため、推薦処理も10秒以下で完了する必要がある。もちろん、処理を高速に行うには、使用するCPUやメモリを高価なものとする

ばよいが、これは機能搭載対象とするテレビの原価コストを上げることになってしまう。本研究では特に300MHzのCPU、256MByteのメモリを搭載した普及型のテレビを対象として、原価コストを上げることなくユーザにとって使い勝手の良いテレビ番組推薦機能を付加可能とするため、図4-1に示すテレビ番組推薦機能を実現する。

《目標》

- (1) CPU負荷300MHz, メモリ量256MByteで10秒以内に推薦処理を完了すること
- (2) 推薦再現率と適合率のバランスをとることができる推薦方式とすること

図4-1 テレビ番組推薦機能の実現目標

4.2 関連技術

テレビ番組を含め、動画だけでなく、広く一般的な情報やアイテムを推薦する方式として、図4-2に示すコンテンツベース方式と協調フィルタリング方式があり[51]、これらを組合わせたハイブリッド方式も提案されている[52]。コンテンツベース方式は、推薦対象のユーザと情報やアイテムにおける属性に基づいて、ユーザの属性に合った情報やアイテムを推薦する方式であり、ユーザ個人の嗜好を用いるため、ネットワークに接続されない環境にも適用可能な方式である[53]。他方、協調フィルタリング方式は、推薦対象とするユーザに対して、好みの類似しているユーザが高評価を与えている情報やアイテムを推薦する方式であり、コンテンツベース方式で取得が困難かつ有効性の判断が難しい属性情報を用いる必要がないため、昨今では広く研究が進められている方式である[54]。これらの方式には上記以外にも一長一短があり、コンテンツベースの推薦方式では、情報やアイテムの属性を利用して、ユーザの好みに合った新たな情報やアイテムを推薦しやすい反面、ユーザが興味を持ったことのない意外な情報やアイテムを推薦することは困難である

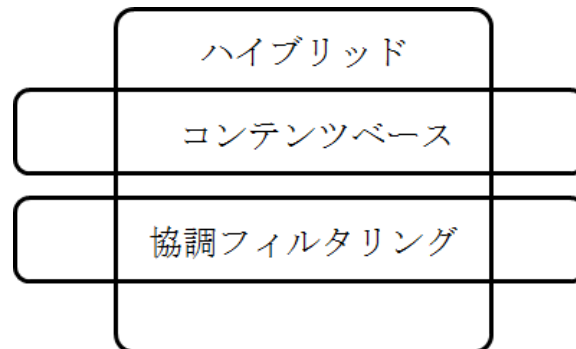


図4-2 テレビ番組推薦機能を実現する技術

と言われている[55]．一方，協調フィルタリング方式では，他のユーザが評価していない新規の情報やアイテムを推薦することが困難である反面，対象ユーザが興味を示したことのない意外なアイテムが推薦されやすいという特徴がある[55]．ハイブリッド方式はそれぞれの方式の長所を活かし，欠点を補う方式として提案されている[52]．以下ではこれらの方式に関して本研究に関連する先行技術について述べる．また，本研究に係る推薦精度，推薦番組数，及び推薦処理負荷に関連した技術について述べる．

コンテンツベース方式による推薦技術に関しては，ユーザのテレビ番組視聴履歴を用いた推薦技術が提案されている[56]．本技術では，番組説明中の出演者名とキーワードに対して「嗜好度」を計算することで，ユーザの嗜好に基づいたテレビ番組の推薦が可能となる．嗜好度の計算（以下，嗜好抽出と呼ぶ）方法に関しては，番組説明中の出演者名とキーワードに加えて，さらに番組タイトルも考慮して計算する方法[57]や番組ジャンルも考慮して計算する方法[58]及び放送チャンネルも考慮して計算する方法[59]が提案されている．これらにより，テレビ番組に対するより広範囲のユーザの嗜好を把握することが可能となる．これらの技術では番組に対するキーワードの抽出及び活用が提案されているが，嗜好抽出におけるキーワードの妥当性は考慮されておらず[60]，実際にはキーワードにはノイズがあるため，精度劣化の

要因となる場合も多い[53][55]. そのため, キーワードに基づく推薦と, 出演者に基づく推薦と, キーワード+出演者に基づく推薦, 及び時間帯+チャンネルに基づく推薦を評価し, 習慣性を予測できるという点で, 時間帯+チャンネルによる推薦手法の有効性を示した成果も報告されている[60]. さらに, 番組タイトルと照合することで, 毎週放送されるシリーズ番組を検出し, 番組に関する情報を予測する技術も提案されている[61]. これにより, テレビ番組の推薦に必要なデータを補完し, より適切な番組を推薦することが可能となる. これらはテレビ番組の視聴操作に基づいて嗜好抽出を行うが, 嗜好抽出には録画等の操作を加味することが重要であるとの指摘もある[55]. 視聴操作のみからの嗜好抽出では, ユーザが「ながら観」している場合や, テレビでは表示されていても実際にはユーザが観ていない場合にも嗜好に反映されてしまう. これに対して, 録画操作はユーザが好む番組に対して明示的に行われるため, 録画操作を用いることで, テレビ番組に対するユーザの嗜好をよりの確に把握することが可能となる.

協調フィルタリング方式による推薦技術に関しては, ネット動画における再生リストの類似性を用いた推薦技術が提案されている[62]. 本提案では, 動画の再生にはユーザの嗜好が多く含まれていることを検証している. このため, テレビ番組に対しても再生操作の活用が嗜好の抽出に有効であると言える. また, YouTube[63]等の動画共有サイトにおいて, 動画投稿者等のユーザのつながりを用いた推薦方式が提案されている[64]. 本技術では, ユーザが閲覧した動画の投稿者の周辺ユーザ(支持者, 非支持者, 友達等)を巡って関連動画を集め, 推薦対象のユーザが閲覧した動画のメタデータから特徴語を生成し, この特徴語をメタデータに含む動画を推薦している. メタデータとしては動画のタイトルや説明文から得られるキーワード及びタグを利用しており, これらにより, 推薦対象のユーザが閲覧した動画に対して関連と意外性のある動画の推薦が可能となる. 同様に, YouTubeにおける動画推薦において, 動画に付与されているタグ情報を応用した推薦方式として, タグ情

報から得られる階層的なキーワードを用いた推薦方式も提案されている[65]. 本方式では各動画に付与されているタグの意味的な包含関係を用いて、抽象的なキーワードを用いて推薦対象の動画に対するユーザの嗜好度を示す推薦スコアが計算される. このため、好みの動画に限らず多様な動画（意外な動画）を推薦可能となる. これらのタグ情報はネット動画に特有の情報ではあるが、タグに限らず動画に付随するキーワードを用いることが可能であると考えられ、テレビ番組に対しても番組説明文のキーワードを用いることで同様の機能を実現できるものと考えられる.

ハイブリッド方式による推薦技術としては、ユーザから入力されるプロフィールに基づくコンテンツベース方式による推薦と協調フィルタリングによる推薦を組合わせたテレビ番組の推薦技術が提案されている[66][67]. 本方式では、視聴したテレビ番組のチャンネル、ジャンル、出演者名、キーワードに加えて、視聴時間、言語、制作国、制作者名、監督名、ディレクター名を使った嗜好抽出方式が提案されている. これにより、ユーザがこれまで見たこともない多様な動画を推薦可能となる. 組合わせるコンテンツベース方式において、テレビ番組におけるジャンル、出演者、ディレクタとキーワードを重視して嗜好抽出を行うと共に、協調フィルタリング方式において、ユーザからのレーティングに録画実績を加算して推薦スコアを計算する方法も提案されている[68]. これによりユーザの嗜好に合った推薦と意外性のある推薦を実現することができる. さらに、ソーシャルネットワークでのタグやレーティング及びコメントを利用した推薦スコアの計算方法も提案されている[69]. この方式でもユーザの好みに合ったテレビ番組だけでなく、意外性のある多様なテレビ番組を推薦可能となる. また、これらの本方式では、コンテンツベース方式の推薦により新規ユーザ及び新規動画の推薦を行い、協調フィルタリング方式の推薦によりユーザが知らなかった多様な動画を推薦可能となる[70].

これらの関連技術の中には、本研究で対象としているように、推薦におけ

る適合率と再現率がトレードオフの関係にあり、目的に応じて重要視する指標を選定する必要があることを述べているものもある。神鷲らは、場合によってはこれらのバランスをとる必要があることを指摘している[55]。また、どの指標をどのようなバランスで重要視するかは、推薦対象やユーザの目的など多くの要因を考慮して決めなければ、ユーザの満足を得られるような推薦システムは設計できないとされている[71]。このため、推薦の質を高めるためのユーザ理解方法とフレームワークも提案されている[72]。このフレームワークはユーザの経験（Human-Recommender Interaction）に基づいて設計されており、これを用いることでユーザにとってどのような推薦が良いかを検討することができる。特に、推薦結果のリストを対象として、その良し悪しを図る指標を検討することが可能となる。また、数々の指標でユーザの満足度を高める推薦結果を検討し、ユーザのニーズに合った推薦を実現することができる。これらは評価の方法論であり、実現方法や技術に関しては触れられていない。一方で、これらの評価軸のすべてを満足するには推薦の種類を多くする必要があるため、コスト的に実現が困難ではあるものの、推薦システムの対象によって評価軸を選定する必要性は高いと考えられる。これを鑑み、テレビ番組の推薦では、最近のトレンドからテレビ視聴時間が増加しえないことと、見たい番組のすべてを知った上での推薦指標に疑問があることから、推薦おける適合率を重視し、これを高める技術が提案されている[53]。逆に、再現率を重視し、これを高める技術も提案されている[59]が、本研究で目標としているように最適な推薦番組数で適合率と再現率のバランスを最適化する技術に関しては提案されていない。

推薦番組数に着目した技術としては、一日当たり5番組と固定としている技術[53][60]や1回の問い合わせに対して1番組を推薦する技術[57]がある。これらに関しては、ユーザ及び時と場合に合わせた推薦の必要性は述べているが、具体的な対策は行われていない[57]。その中でも可能な限り無効な番組を推薦しないよう、シリーズ番組を推薦しない方式も提案されており[61]、

実際にテレビにおける自動録画機能に適用されている。また、テレビ番組等の動画推薦を対象としたものではないが、利用者からアイテムへの好き嫌い評価するだけでなく、既知かどうかの情報も得ることで、ユーザが未知であると予測されるアイテムを推薦する技術が提案されている[73]。これにより有効な推薦番組を自動録画あるいは推薦表示することが可能となるが、推薦における適合率と再現率のバランスを良くする考えはない。テレビ番組や動画に限らず推薦アイテム数を制限する技術そのものが少ないと指摘されている[55]が、推薦対象のアイテムに対するユーザの嗜好度を示す推薦スコアの高い順に上位N個の推薦を行う技術も提案されている[55][74]。このように推薦数を制限することで限られた表示画面で効果的に動画を推薦できる[68]。また、推薦スコアにおける上位N個の推薦結果に対して、推薦アイテムを関連度の異なるアイテムと入れ替える技術も提案されている[75]。これにより推薦結果に対する関連度の偏りをなくし、多様性のある推薦を実現することが可能となる。テレビ番組の推薦に関しても、通常はスコアの高いN個のテレビ番組が推薦されることを示唆した上で、習慣的に視聴しているテレビ番組を推薦しないように配慮したシステムも提案されている[76]。これにより、限られた推薦番組数の中で、意外性のある番組を推薦することが可能となる。これらは多様性や意外性のある推薦に注力しているため、必ずしも推薦における適合率と再現率のバランスが良くするものではない。また、ネット動画の推薦に関して、スコアの高いN個の動画を推薦することをベースとして、ユーザの視聴履歴やタイトル及び説明文に基づく推薦結果と、ネット上での他のユーザとのレーティングやコメント及び共有等に基づく推薦結果が一致した動画を推薦することで、推薦アイテム数を決定する技術も提案されている[77]。本技術により、限られた推薦数で、ユーザがこれまでに視聴した動画に対する関連性と多様性を確保した推薦結果を得ることが可能となる。しかしながら、これらの方式もやはり本研究で対象とする推薦における適合率と再現率のバランスを良くするものではない。

意外性や多様性のある推薦を行うための技術としては、コンテンツベースと協調フィルタリングのハイブリッド推薦により意外性のある推薦を行うことが提案されている[78]。本技術により、ユーザの気づかなかった番組を推薦できる。また、ユーザが認識しているアイテムクラスタの距離を考慮して、それらのクラスタ間のアイテムを推薦する方式も提案されている[79]。これにより、意外性のある推薦を行うことが困難なコンテンツベースの推薦において、意図的に意外性のある推薦を行うことが可能となる。さらに、コンテンツベース推薦と協調フィルタリング推薦に加えて、試し推薦を行う技術が提案されている[80]。本技術によりユーザはこれまで気づかなかった新たなアイテムを発見することができ、ユーザの嗜好空間を広げることができるため有効な技術であるといえる。これらの技術は、いずれも情報推薦の活性化を実現する非常に有用な技術である。一方で、テレビでは一度好ましくない結果が表示された場合にクレームがつく可能性や二度と使用してもらえなくなる可能性もあり注意を要する。このため、これらの技術は、ユーザの好む番組に限られた表示領域に表示する或いは限られたハードディスクに自動録画する観点には向かない可能性も高い。

処理負荷や必要計算リソースに関しては、コンテンツベース方式による推薦では、ユーザの好むキーワードに基づく推薦を行うために、全番組のタイトル及び説明文に対して形態素解析を行い、名詞や人名を判断する必要がある。これには処理負荷がかかり、多くの計算リソースが必要となる。このため、ユーザインタフェースに影響を与えないよう、短時間でテレビ番組の推薦を行うための処理負荷及び必要計算リソースの削減は今後の課題としてあげられている[64]。これに対して、組込み機器でテレビ番組をリアルタイムに推薦する技術が提案されている[81]。本技術では、計算量が多いキーワード抽出を事前に処理することで推薦処理の負荷低減による高速化を実現しており、具体的な処理負荷及び必要計算リソースに関する記載はないものの、実際の製品に適用する上では有効な技術である考えられる。また、動画のタ

グ情報をカテゴリーに分類し、カテゴリーごとにユーザの嗜好を計算して、ユーザの好むカテゴリーの動画を推薦する方式が提案されている[82]。これにより、ユーザの嗜好をカテゴリーごとに判断できるため、推薦における処理負荷を軽減し、必要計算リソースを低減することができる。本技術では、動画の重みを他のユーザが設定しており、その重みに基づいてユーザの好みを計算しているため、結果的に嗜好の類似しているユーザが高評価を与えた動画を推薦していると考えられ、広義の意味では協調フィルタリング方式による推薦であるとも考えることができる。協調フィルタリング方式による推薦に関しては、ユーザによって選択されたアイテムを I_1 として、 I_1 を選択した他のユーザの既選択アイテムの中から I_1 と類似なアイテム I_2 を推薦する技術も提案されている[83]。本技術では、全ユーザの類似度を求める必要はなく、同一アイテムを選択した人のみを対象とできる。このため、これらの方式では処理負荷を軽減し、必要計算リソースを低減できると考えられるが、具体的な処理負荷及び必要計算リソースに関しては明記されていない。ただし、同様な考え方でユーザクラスタリングにより類似ユーザの相関計算を削減することで、高速に協調フィルタリングを行う方式も提案されており[84]、本方式では、300MHzのCPU、512MByteのメモリでユーザ1人及び1アイテムあたり、13ミリ秒程度で処理可能となると考えられる。この場合、日本の地上波デジタルテレビ放送で提供される1週間分（約3000番組）の処理には39秒程度必要となり、本研究で目標としている300MHzのCPU、256MByteのメモリで10秒以内に推薦処理を完了することは困難であると考えられる。このため、コンテンツベース方式と協調フィルタリング方式からなるハイブリッド方式も協調フィルタリングよりさらに処理時間がかかり、本研究で目標としている10秒以内に推薦処理を完了することは困難であると考えられる。

以上をまとめると表4-1及び表4-2の通りとなる。表4-2に示す通り、関連技術では、本研究で目標としている「推薦における適合率と再現率のバランスをとる」ことはできないため、これらによるテレビ番組推薦機能は、ユー

ずにとっての使い勝手が良くないものになってしまう。また、処理負荷及び必要計算リソースについても、嗜好抽出に用いる番組属性としてのキーワード抽出に処理負荷を要し、必要計算リソースが多くなると考えられ、処理負荷及び必要計算リソースの削減は今後の課題とされている。特に、高速に協調フィルタリングを行う方式でも300MHzのCPU、512MByteのメモリでユーザ1人及び1アイテムあたり、13ミリ秒程度の処理が必要となる。このため、日本の地上波デジタルテレビ放送で提供される1週間分（約3000番組）の処理には39秒程度必要となると考えられ、関連技術の協調フィルタリング及びハイブリッド方式では本研究において目標としている300MHzのCPU、256MByteのメモリで10秒以内に推薦処理を完了することはできないと考えられる。したがって、これらによってテレビにおける番組推薦機能を実現する場合には、CPUやメモリを高価なものにする等、コストを犠牲にするか、ユーザにとっての使い勝手を犠牲にする必要がある。

表4-1 推薦方式に関する関連技術のまとめ

推薦方式	特徴
コンテンツベース方式	<ul style="list-style-type: none"> ・ネットワークに接続されない環境にも適用可能 ・ユーザの好みに合った新たな情報やアイテムを推薦可能 ・適合率と再現率は良いが、意外な情報やアイテムを推薦することは困難
協調フィルタリング方式	<ul style="list-style-type: none"> ・ネットワーク接続が必要 ・新規の情報やアイテムを推薦をすることが困難 ・取得が困難かつ有効性の判断が難しい属性情報を用いる必要がない ・対象ユーザが興味を示したことのない意外なアイテムを推薦可能
ハイブリッド方式	<ul style="list-style-type: none"> ・ネットワーク接続が必要 ・ユーザの好みに合った新たな情報やアイテムを推薦可能 ・対象ユーザが興味を示したことのない意外なアイテムを推薦可能

表4-2 適合率と再現率のバランス及び処理負荷に関する

関連技術のまとめ

項目	現状
適合率と再現率のバランス	推薦アイテム数は固定（推薦スコアの高い順に上位N個）であり，適合率と再現率のバランスを最適化する技術に関しては提案されていない
処理負荷	<ul style="list-style-type: none"> ・コンテンツベース方式では嗜好抽出に用いる番組属性としてのキーワードの抽出に処理負荷がかかると考えられ，処理負荷の削減は今後の課題として挙げられている ・高速協調フィルタリング方式では300MHzのCPU，512Mbyteのメモリでユーザ1人及び1アイテムあたり，13ミリ秒程度必要（日本の1週間分の処理には39秒程度必要）であり，関連技術の協調フィルタリング方式及びハイブリッド方式により本研究で目標としている「300MHzのCPU，256Mbyteのメモリで10秒以内に推薦処理を完了する」ことは困難

一方，テレビ番組推薦では推薦精度が高いことが要求されることが述べられており[53]，番組が毎回変化しているため，この条件下で高精度化を期待できるコンテンツベース方式が良いとされている[76]．また，テレビはネットワークに接続されていないケースも多く，新規番組を受信することも多い．このため，この観点でもネットワーク接続を前提とせず，新規アイテムを推薦しやすいコンテンツベース方式が有利である．そこで，本研究ではコンテンツベース方式を対象とし，関連技術では成し得ていない本研究の目標である「CPU負荷300MHz，メモリ量256MByteで10秒以内に推薦処理を完了すると共に，推薦再現率と適合率のバランスをとることができる推薦方式」を実現する．これによりテレビの原価コストを上げることなく，ユーザにとって使い勝手の良いテレビ番組推薦機能を普及型のテレビに付加可能とする．

4.3 提案方式

前述の目標を実現するため、本研究では次のステップからなるアプローチを取った(図4-3)。

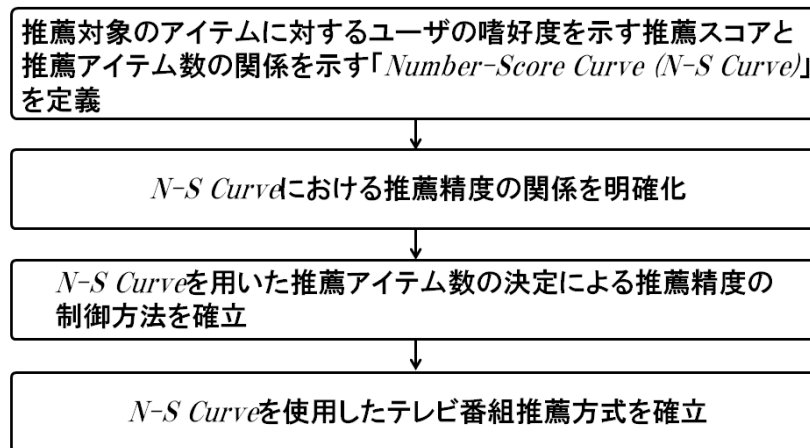


図4-3 提案方式確立に向けたアプローチ

ステップ1：推薦対象のアイテムに対するユーザの嗜好度を示す推薦スコアと推薦アイテム数の関係を示す“Number-Score Curve (N-S Curve)”を定義する

ステップ2：N-S Curveにおける推薦精度の関係を明確化する

ステップ3：N-S Curveを用いて、推薦スコアと推薦精度に依存する推薦アイテム数を制御することで、推薦精度を向上させる方法を確立する

ステップ4：N-S Curveを使用したコンテンツベース方式に基づくテレビ番組推薦方式を確立する

以下では、これらについて説明する。

4.3.1 Number-Score Curve (N-S Curve)

推薦システムでは、通常、アイテム毎に推薦スコアを計算する。推薦スコアは、テレビ番組等の推薦対象のアイテムに対するユーザの嗜好度あるいはユーザの好みに対する合致度を示すものであり、ユーザの好みの属性に合致しているアイテムほど推薦スコアが高くなる。具体的には、例えば、あるユーザが「芸能人A」を好み、ドラマが好きな場合には、芸能人Aが出演しているドラマの推薦スコアは高くなる(図4-4)。また、芸能人Aが出演しているドラマ以外の番組のスコアは中程度のものとなり、芸能人Aが出演していないドラマ以外の番組の推薦スコアは低くなる。

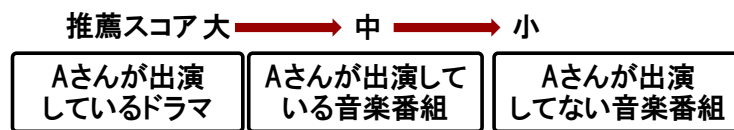


図4-4 推薦スコア

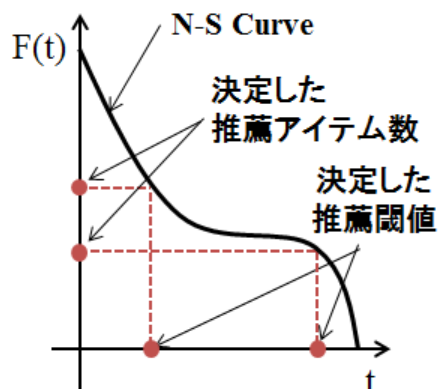


図4-5 N-S Curveの例

これに対して、ある推薦スコア以上のアイテムを推薦することを想定し、ある推薦スコア以上のアイテム数を示した曲線をN-S Curveとして定義する。すなわち、N-S Curveは、アイテム k の推薦スコアを S_k とすると、以下の関数で与えられる。

$$F(t) = \left| \{ k \mid S_k \geq t \} \right|, \quad 0 \leq t < \infty \quad (4.1)$$

この定義より、N-S Curveは推薦スコアが t 以上のアイテム数を表し、 t の減少関数である(図4-5)。よって、 t との $F(t)$ 関係は次式の通りとなる。

$$\text{If } t \uparrow \text{ then } F(t) \downarrow \quad (4.2)$$

$$\text{If } t \downarrow \text{ then } F(t) \uparrow \quad (4.3)$$

この定義では、推薦スコアが t 以上のアイテムを推薦アイテムとしたとき、 $F(t)$ は推薦アイテムの総数を示している。以下では $F(t)$ における t を推薦閾値と呼び、推薦閾値 t における $F(t)$ の値が推薦アイテム数と呼ぶ。

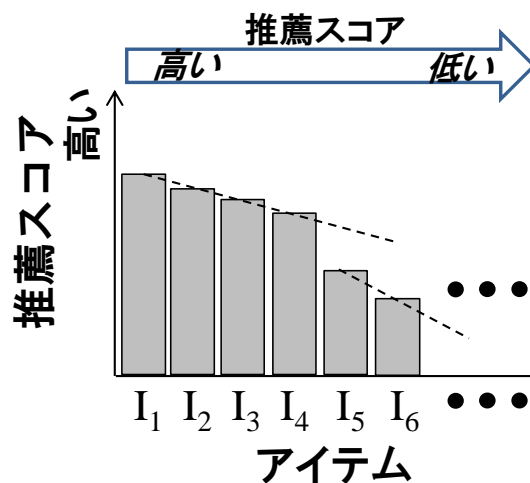


図4-6 推薦スコアでソートしたアイテム

別の観点では、N-S Curveは、推薦スコアの高い順に推薦アイテムを並べているといえる。図4-6はこの観点を示しており、横軸にアイテムを並べ、縦軸に推薦スコアをプロットしたグラフである。このグラフではすべてのアイテムを推薦スコアの高い順に並べている。このグラフからもわかるとおり、すべてのアイテムを推薦スコアの高い順に並べたとき、推薦スコアの減衰が線形になることは稀である。そのため、N-S Curveは、連続値を持つ近似曲線とみなすことで、一つ以上の変曲点を持つこととなる(図4-7)。

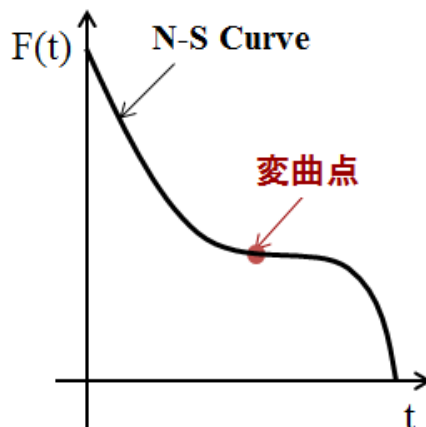


図4-7 N-S Curveにおける変曲点

4.3.2 N-S Curveにおける推薦精度の関係

推薦システムにおける推薦精度を測る指標としては、適合率と再現率が知られている[55]。適合率は推薦結果の正しさを示す指標であり、再現率は推薦の網羅性を示す指標である。推薦システムでは、これらを総合的に高めることが必要となる。適合率と再現率は次の通り表される。

$$Pr = n(D \cap R) / n(R) \quad (4.4)$$

$$Re = n(D \cap R) / n(D) \quad (4.5)$$

ここで、 Pr と Re は、それぞれ適合率と再現率を示す。また、 D と R は、それぞれユーザが望むアイテムの集合と推薦されたアイテムの集合を示している(図4-8)。

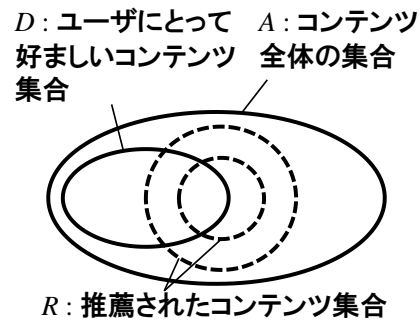


図4-8 適合率と再現率

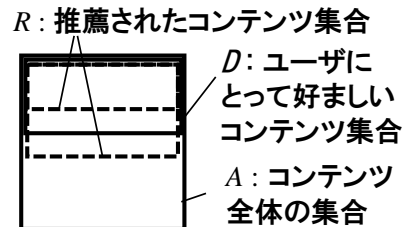


図4-9 推薦アイテム数と適合率及び再現率の関係

N-S Curveおよび図4-6のようにアイテムを推薦スコアの高い順に並べた場合、推薦アイテム数を増加させると、図4-9に示すとおり、ユーザが望むアイテムがより多く推薦される確率が高くなるため、再現率が良くなり、適合率が悪くなると考えられる。逆に、推薦アイテム数を減少させると、推薦数が少なくなる一方で、ユーザが望むアイテムが推薦される確率が低くなるため、

適合率は良くなるが，再現率は悪くなると考えられる．この関係を示すと以下の通りとなる．

$$\text{If } F(t) \searrow \text{ then } Re \searrow \text{ and } Pr \nearrow \quad (4.6)$$

$$\text{If } F(t) \nearrow \text{ then } Re \nearrow \text{ and } Pr \searrow \quad (4.7)$$

したがって，N-S Curveを用いることで，推薦スコアに対する推薦閾値と推薦アイテム数及び推薦精度の間には以下の関係を導き出すことができる．

- (1) 推薦閾値を大きくした場合，推薦アイテム数は減り，推薦適合率は良くなる．一方で，推薦再現率は悪化する
- (2) 推薦閾値を小さくした場合，推薦アイテム数は増加し，推薦再現は良くなる．一方で，推薦適合率は悪化する

これらは以下の通り表現できる．

$$\text{If } t \nearrow \text{ then } F(t) \searrow, Re \searrow \text{ and } Pr \nearrow \quad (4.8)$$

$$\text{If } t \searrow \text{ then } F(t) \nearrow, Re \nearrow \text{ and } Pr \searrow \quad (4.9)$$

これらの関係から，N-S Curveを用いて推薦スコアに対する推薦閾値を調整することで，間接的に推薦アイテム数をコントロールし，結果的に推薦における適合率と再現率のバランスを向上させることが可能となる．

4.3.3 推薦アイテム数決定方法

本節では，推薦における推薦アイテム数を決定するための方法を提案する．

4.3.1節で述べたとおり，通常では，N-S Curveは一つ以上の変曲点を持つ

(図4-7). そこで、提案方式では、N-S Curveにおける最大変曲点を見つけ出し、この最大変曲点に対応する推薦閾値を、推薦アイテム数を決定するための推薦閾値（推薦アイテム数決定推薦閾値）と定める．すなわち、この推薦アイテム数決定推薦閾値以上の推薦スコアを持つアイテムを推薦アイテムとすることで、推薦アイテム数を決定する．ここで、推薦アイテム数決定推薦閾値すなわち最大変曲点に対する推薦閾値は、N-S Curveにおいて推薦アイテム数に変化がない t の最大区間における中点である．これは以下により求めることができる．

- (1) N を全アイテム数とし、 $i=\{1,2,\dots,N-1\}$, $n=\{1,2,\dots,N-1\}$ として、次式を満たす2つの推薦閾値 t_i と t_{i+1} を求める

$$\begin{cases} t_i < t_{i+1} \\ F(t_i) - F(t_{i+1}) = 0 \\ t_{i+1} - t_i = \max(t_{n+1} - t_n) \end{cases} \quad (4.10)$$

- (2) 次式のとおり、 t_i と t_{i+1} の平均 T を推薦アイテム数決定推薦閾値とする

$$T = \frac{t_i + t_{i+1}}{2} \quad (4.11)$$

実際の処理では推薦スコアは離散値となるため、N-S Curveは例えば図4-10のようになる．この場合、推薦アイテム数に変化がない t の最大区間は①の区間であり、その中間点である T を推薦アイテム数決定推薦閾値として設定する．この最大区間は、意味的には図4-11に示す通り推薦スコアでソートしたアイテム空間における推薦スコアの最大ギャップ(最大推薦スコアギャップ)に対応している．このため、N-S Curveに基づく上述した処理により、この最大区間を求めるだけで、最大推薦スコアギャップを求めることができる．また、この最大区間の中点を推薦アイテム数決定推薦閾値とすることで、最大

推薦スコアギャップの中間位置に推薦アイテム数決定推薦閾値を決定することができる(図4-11).

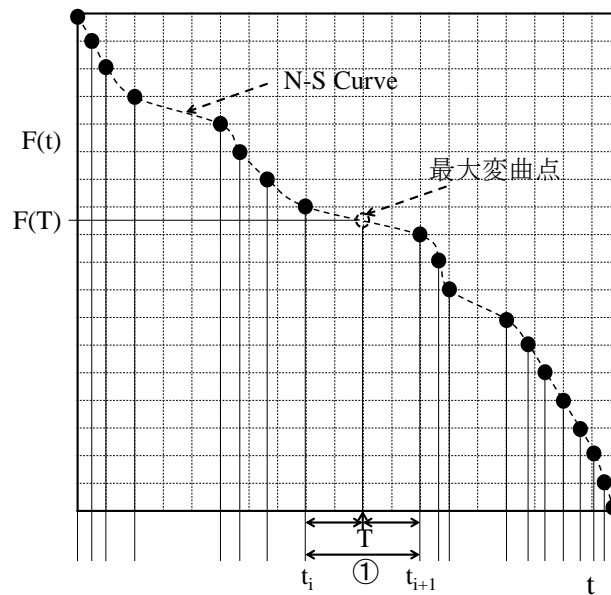


図4-10 推薦アイテム数決定推薦閾値の選定

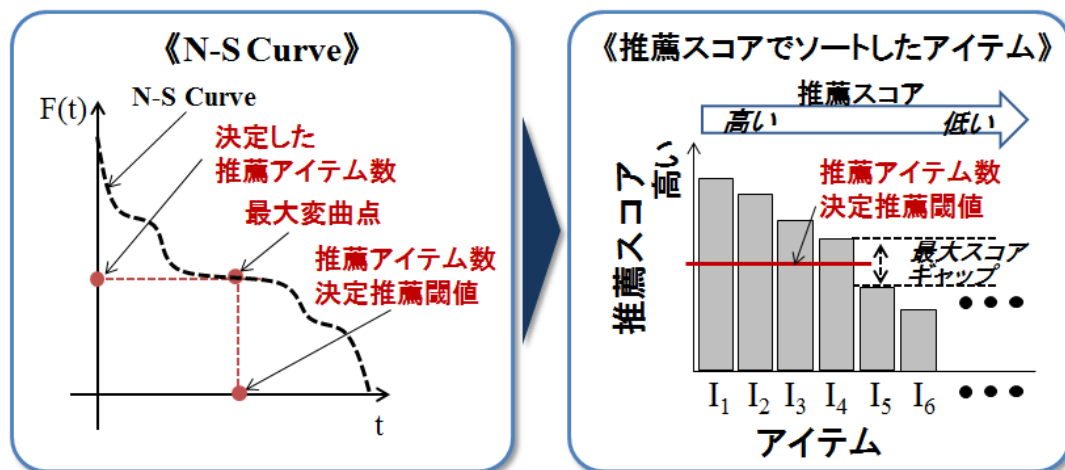


図4-11 推薦アイテム数の決定

通常，最大推薦スコアギャップの中間位置に推薦アイテム数決定推薦閾値を決定しようとした場合，上述したとおり，すべてのアイテムを推薦スコアでソートして最大推薦スコアギャップを求める必要がある．一方で，N-S Curveはアイテムをソートする必要が無く，予め定めたスコア範囲に属するアイテム数をカウントするだけで生成できる．このため，上述のとおりN-S Curveを用いて推薦アイテム数に変化がない t の最大区間を求めることで，処理負荷を軽減し，必要計算リソースを低減して推薦アイテム数決定推薦閾値を設定し，推薦アイテム数を決定することが可能となる．さらに，本方式では，図4-12に示すとおり，推薦を実施する日時及びユーザ毎に推薦アイテム数を決定することが可能となる．

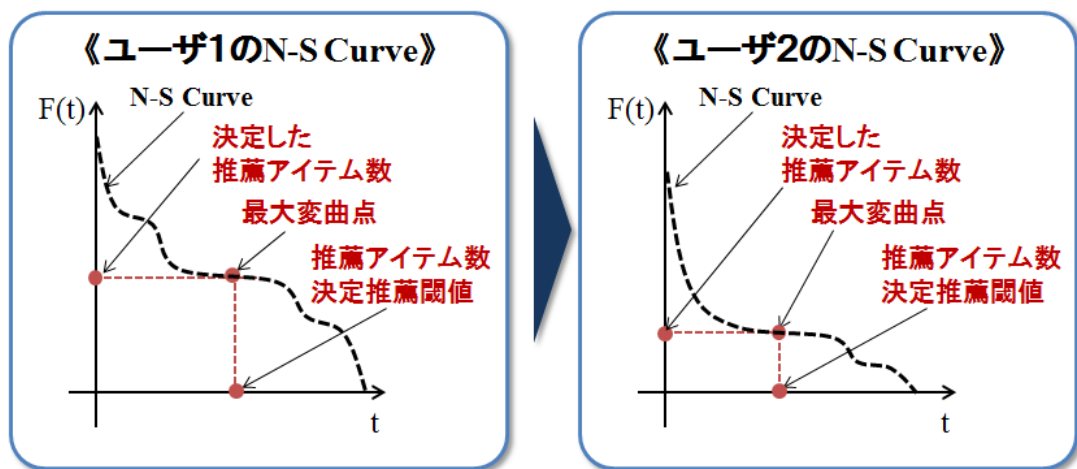


図4-12 ユーザごとの推薦アイテム数の決定

4.4 テレビ番組推薦への適用

本提案方式を用いて，普及型の録画機能付きテレビ向けのテレビ番組推薦機能を実現した．

4.2節で述べたとおり，推薦方式としては協調フィルタリング方式と，コン

テンツベース方式，及びこれらを組合わせたハイブリッド方式が知られている．一方で，テレビ番組推薦では推薦精度が高いことと，放送される番組が常に変化していることを考慮する必要がある．また，テレビはネットワークに接続されていないケースも多い．さらに，新規番組が放送されることも考慮する必要がある．そこで，本研究ではこれらに有利とされているコンテンツベース方式適用することとした．

ここで，コンテンツベースの方式では，ユーザの嗜好を解析するために用いるアイテムの属性と嗜好の抽出方法が処理負荷及び推薦精度を考慮する上で重要となる[85]．そのため，本実現においては，ユーザの嗜好を的確にあらわす必要最低限のテレビ番組の属性と操作を定義した．具体的には，アイテムであるテレビ番組の属性として，EPG(Electronic Program Guide)から取得できる，(A1) 番組タイトル(Title)，(A2) 番組内容のキーワード(Keywords)，(A3) 番組ジャンル(Genre)，(A4) 放送チャンネル(Channel)，及び(A5) 放送時間(Time)を定義した．また，番組内容のキーワードとして，番組ジャンルに適応したキーワードを表4-3の通り定義した．これらはEPGデータにおけるタグ処理のみで実現することができ，4.2節で述べた処理負荷及び必要計算リソースを増やす要因となる形態素解析を行う必要がない．このため，テレビ番組推薦における処理負荷を軽減し，必要計算リソースを低減することが可能となる．

表4-3 ジャンル適応キーワード

ジャンル	キーワード
ドラマ, 映画, 音楽, バラエティ	出演者及びアーティストの名前
スポーツ	チームの名前
その他	出演者の名前

また，テレビ番組の操作としては，(O1) 再生 (Playback)，(O2) 予約 (Reserve)，(O3) 検索 (Search)，及び(O4) 録画 (Recording) を定義した．
4.2節で述べたとおり，従来はテレビ番組の視聴のみを対象としていたのに対

表4-4 嗜好抽出に用いる属性と重み


A_i	属性 (A)	重み (W_a)
A_1	タイトル	
A_2	キーワード	
A_3	ジャンル	
A_4	チャンネル	
A_5	放送時間	

表4-5 嗜好抽出に用いるテレビ操作と重み

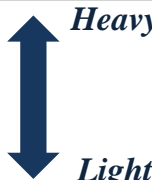
O_j	操作 (O)	重み (W_o)
O_1	再生	
O_2	予約	
O_3	検索	
O_4	録画	

表4-6 属性に対する重みの例

A_i	属性 (A)	重み (W_a)
A_1	タイトル	2.0
A_2	キーワード	1.4
A_3	ジャンル	1.2
A_4	チャンネル	0.1
A_5	放送時間	0.1

し、これらの操作を重要視することで、単なる「ながら観」によるノイズを排除し、よりユーザの嗜好を的確に抽出することが可能となる。

ユーザの嗜好は、これらの属性と操作に基づく統計計算によって求める。特に、これらの属性において、重みを表4-4のとおり定めた。また、操作については、重みを表4-5のとおり定めた。これらの重みに対して、推薦計算の対象となるテレビ番組に出現する属性数に基づいて正規化し、実際の重み値を設定した。表4-6及び表4-7は、それぞれ、属性の重み値と操作の重み値の一例を示している。

表4-7 操作に対する重みの例

O_j	操作 (O)	重み (W_o)
O_1	再生	1.4
O_2	予約	1.3
O_3	検索	1.2
O_4	録画	1.1

次に、これらの属性と操作に基づいてユーザ嗜好データベースを生成する(図4-13)。本データベースは、ユーザの嗜好に関する情報を格納したデータベースである。本データベースは、先に示したそれぞれの属性に対するテーブルを持ち、それぞれのテーブルは、さらに属性のエントリと統計値を持つ。

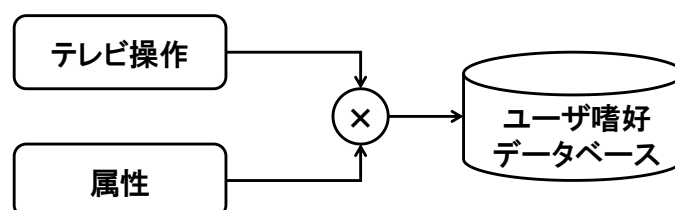


図4-13 ユーザ嗜好データベース

例えば，ユーザ嗜好データベースにおけるキーワードテーブルは，表4-8に示すとおりエントリである出演者の名前と統計値からなるレコードを持つ．同様に，ジャンルテーブルは，表4-9に示すとおりエントリであるジャンル名と統計値からなるレコードを持つ．

表4-8 キーワードテーブルの例

キーワード	統計値
松田聖美	5.0
小泉明日子	4.0
河合直美	3.0

表4-9 ジャンルテーブルの例

ジャンル	統計値
ドラマ	5.0
音楽	4.0
ニュース	3.0

これらの統計値は，テレビ番組が操作される度に更新され，操作された番組のEPGに含まれる属性におけるエントリ毎に，その属性の重み値と操作の重み値を掛け合わせて，加算していくことで計算される．すなわち，例えば，キーワードテーブルの場合，EPGに出演者Aが含まれている番組を録画した場合には，出演者Aがエントリとなり，先の表4-6及び表4-7を用いて，このエントリに対するこれまでの統計値に，「録画した」という操作の重み値と，「キーワード」の重み値を掛け合わせた「 $1.1 * 1.4$ 」を加算する．

表4-10を用いて詳細に説明する．表4-10は，あるユーザによるテレビ操作ログ($L_1 \sim L_n$)と，そのログに含まれる属性を示している．例えば#1のレコードはタイトル t_1 ，ジャンル g_1 ，キーワード k_1 ，チャンネル c_1 ，放送時間 b_1 のテレビ番組に対してユーザが p_1 の操作を行ったことを示している．

表4-10 テレビ操作ログ

#	タイトル	ジャンル	キーワード	チャンネル	放送時間	操作
1	t_1	g_1	k_1	c_1	b_1	p_1
...
n	t_n	g_n	k_n	c_n	b_n	p_n

ユーザ嗜好データベースは，このようなテレビ操作ログを用いて生成され，それぞれの属性のエントリに対する統計値は，次式によって算出される．

$$St(e) = Wa(A_i) \sum_{O: \text{Operation for } e \text{ in all the logs}} Wo(O) \quad (\text{For } e \in A_i) \quad (4.12)$$

ここで， $St(e)$ は，各属性に含まれるエントリ(e)に対する統計値を示し， Wa と Wo は，それぞれ属性と操作の重み値を示している．また， O は，すべてのログ(L_k)に含まれる属性のエントリ(e)に対するすべての操作 O_j を示している．したがって，例えば，ユーザが出演者Aが出演しているテレビ番組を録画して再生した場合には，出演者Aというキーワードのエントリに対する統計値は，録画と再生の重み値と，キーワードに対する属性の重み値を掛けたものとなる．すなわち，表4-6及び表4-7で示した重み値の場合，この統計値は次式によって計算できる．

$$St(\text{“出演者A”}) = 1.4 * (1.1 + 1.4) = 3.5 \quad (4.13)$$

推薦対象のテレビ番組に対する推薦スコアは、上記の通り生成したユーザ嗜好データベースに格納された属性のエントリに対する統計値と、当該推薦対象のテレビ番組の属性のエントリ間の一致度に基づいて計算される。具体的には、推薦対象のテレビ番組に含まれている属性のエントリをEPGから取得する。このとき、キーワードについては、番組のジャンルに基づいて表4-3にしたがって取得する。そして、取得した属性のエントリが、ユーザ嗜好データベースに存在していた場合には、ユーザ嗜好データベースに存在する各エントリの統計値を加算することで、次式によって推薦スコアを計算する。

$$Sc(P) = \sum_{Q_p} St(Q_p) \quad (4.14)$$

ここで、 $Sc(P)$ は、推薦対象のテレビ番組(P)の推薦スコアを意味し、 Q は、テレビ番組(P)のEPGから取得した属性のエントリの内、ユーザ嗜好データベース中に存在するすべてのエントリ(e)を意味している。また、 Q_p はエントリ(e)の統計値を示している。

上記の方法ですべての推薦対象のテレビ番組に対して推薦スコアを計算し、4.3.3節で述べたとおり、N-S Curveを用いて推薦アイテム数決定推薦閾値を設定することで、推薦すべきテレビ番組を決定して表示する(図4-14)。

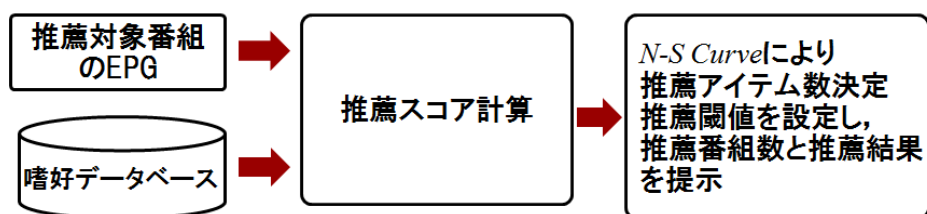


図4-14 推薦番組の決定

表 4-11 テレビ操作ログの例

L_k	Title (A_1)	Genre (A_2)	Keywords (A_3)	Channel (A_4)	Time (A_5)	Operation (O)
L_1	“Ta”	Drama	“Ka”	1	6	Reserve
L_2	“Tb”	Variety	“Kc”	2	5	Reserve
L_3	“Ta”	Drama	“Ka”	1	6	Playback
L_4	“Tb”	Variety	“Kc”	2	5	Playback
L_5	“Tc”	Music	“Kc”	3	6	Search
L_6	“Tc”	Music	“Kc”	3	6	Record

例えば，表 4-11 で示したテレビ番組の操作ログの場合，ユーザ嗜好データベースにおけるテーブルは，表 4-12 乃至表 4-16 の通り計算される．ここで，表 4-11 及び表 4-12 乃至表 4-16 における放送時間 (Time) のエントリは表 4-17 の放送時間帯として表現されるものとしている．また，属性および操作の重み値は，表 4-6 及び表 4-7 で示したものを使用している．これらの例の場合において，表 4-18 で示したテレビ番組の推薦スコアを同じく表 4-18 に示した．

表 4-12 ユーザ嗜好データベースにおけるタイトルテーブル

Title	Statistic	Calculation
“Ta”	5.40	$2.0 * (1.3 + 1.4)$
“Tb”	5.40	$2.0 * (1.3 + 1.4)$
“Tc”	4.60	$2.0 * (1.2 + 1.1)$

表 4-13 ユーザ嗜好データベースにおけるジャンルテーブル

Genre	Statistic	Calculation
Drama	3.24	$1.2 * (1.3 + 1.4)$
Variety	3.24	$1.2 * (1.3 + 1.4)$
Music	2.76	$1.2 * (1.2 + 1.1)$

表 4-14 ユーザ嗜好データベースにおけるキーワードテーブル

Keyword	Statistic	Calculation
“Ka”	3.78	$1.4 * (1.3 + 1.4)$
“Kc”	7.00	$1.4 * (1.3 + 1.4 + 1.2 + 1.1)$

表 4-15 ユーザ嗜好データベースにおけるチャンネルテーブル

Channel	Statistic	Calculation
1	0.27	$0.1 * (1.3 + 1.4)$
2	0.27	$0.1 * (1.3 + 1.4)$
3	0.23	$0.1 * (1.2 + 1.1)$

表 4-16 ユーザ嗜好データベースにおける放送時間テーブル

Time	Statistic	Calculation
5	0.27	$0.1 * (1.3 + 1.4)$
6	0.50	$0.1 * (1.3 + 1.4 + 1.2 + 1.1)$

表 4-17 放送時間 (Time) の表現

Time range	Representation
0:00 - 4:00	1
4:00 - 8:00	2
8:00 - 12:00	3
12:00 - 16:00	4
16:00 - 20:00	5
20:00 - 24:00	6

表 4-18 推薦対象の番組と推薦スコアの計算結果

Ti: Title, Ge: Genre, Ke: Keyword, Ch: Channel, Tm: Time, Sc: Score
 S: Sports, N: News, R: Report, V: Variety, D: Drama, M: Music
 S1: St(α) in Title Table, S2: St(α) in Genre Table, S3: St(α) in Keyword Table,
 S4: St(α) in Channel Table, S5: St(α) in Time Table

Ti	Ge	Ke	Ch	Tm	Statistics					Sc
					S1	S2	S3	S4	S5	
“Te”	S	“Kd”	1	1	0	0	0	0.27	0	0.27
“Tf”	N	“Kf”	2	1	0	0	0	0.27	0	0.27
“Tg”	R	“Kf”	3	2	0	0	0	0.23	0	0.23
“Th”	N	“Kg”	4	2	0	0	0	0	0	0
“Tj”	V	“Kb”	2	3	0	3.24	0	0.27	0	3.51
“Tk”	M	“Kb”	3	3	0	2.76	0	0.23	0	2.99
“Tl”	M	“Kc”	1	4	0	2.76	7.00	0.27	0	10.03
“Tn”	D	“Ka”	4	4	0	3.24	3.78	0	0	7.02
“To”	D	“Kb”	3	5	0	3.24	0	0.23	0.27	3.74
“Tp”	D	“Kc”	1	5	0	3.24	7.00	0.27	0.27	10.78
“Tq”	V	“Ka”	2	6	0	3.24	3.78	0.27	0.50	7.79
“Tr”	D	“Ka”	4	6	0	3.24	3.78	0	0.50	7.52

表4-18から, この例におけるN-S Curveは図4-15の通りとなる. したがって, 4.3.3節で述べたとおり, 最大変曲点に基づいて推薦アイテム数に変化がない最大区間(3.74,7.02]における中点 $T=5.38$ を推薦アイテム数決定推薦閾値に設定することができる. これにより, 推薦番組数を5とし, 推薦アイテム数決定推薦閾値($T=5.38$)よりも大きな推薦スコアを持つ表4-19のテレビ番組を推薦結果すなわち推薦表示すべきテレビ番組として決定することが可能となる.

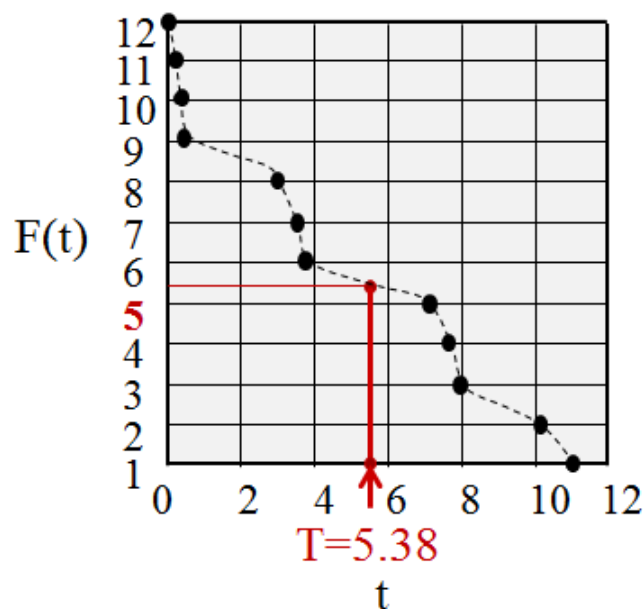


図 4-15 推薦計算例における N-S Curve

表 4-19 推薦表示すべきテレビ番組の決定結果

Title	Genre	Keywords	Channel	Time	Score
"Tl"	Music	"Kc"	1	4	10.03
"Tn"	Drama	"Ka"	4	4	7.02
"Tp"	Drama	"Kc"	1	5	10.78
"Tq"	Variety	"Ka"	2	6	7.79
"Tr"	Drama	"Ka"	4	6	7.52

実際の実現においては、N-S Curveにおける3つの変曲点に基づいて、3つの推薦アイテム数決定推薦閾値を設定し、アイテムの推薦度を決定した。3つの推薦アイテム数決定推薦閾値の設定方法は、4.3.3節で示した方法を拡張し、推薦アイテム数に変化がない t の区間を大きい順に3区間まで求め、これらの区間における中点をそれぞれ3つの推薦アイテム数決定推薦閾値とすることができる。そして、図4-16の通り、これらの推薦アイテム数決定推薦閾値の高い順に推薦スコアの領域を決定し、これらの領域に対して、それぞれ推薦度を高、中、低と決定することで、推薦番組における推薦度を決定する。

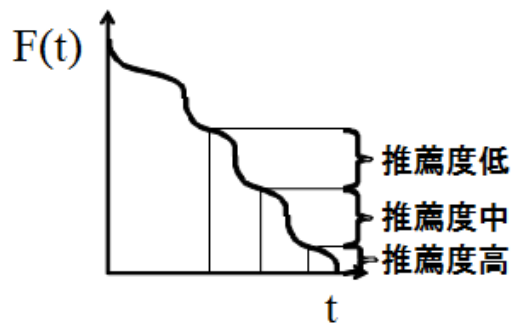


図 4-16 推薦度の決定方法

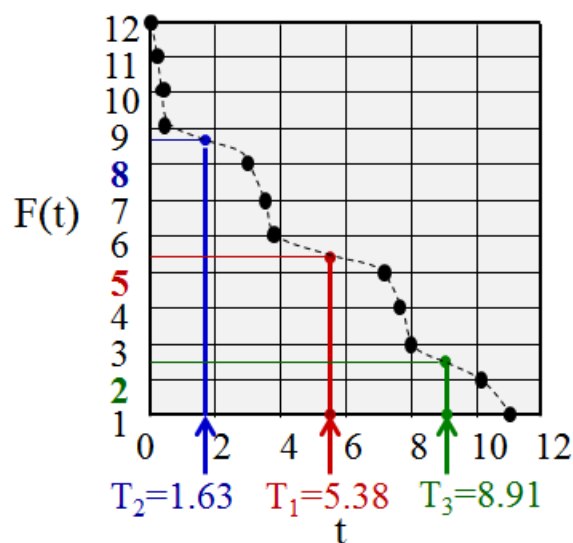


図 4-17 推薦計算例における推薦度の決定

表4-18で示した推薦対象番組及び推薦スコアの計算結果例では，推薦アイテム数に変化がない t の区間を大きい順に3区間まで求めると， $(3.74, 7.02]$ ， $(0.27, 2.99]$ ， $(7.79, 10.03]$ となる．よって，これらの区間における中点 $T_1=5.38$ ， $T_2=1.63$ ， $T_3=8.91$ をそれぞれ3つの推薦アイテム数決定推薦閾値とすることができる(図4-17)．そして，これら3つの推薦アイテム数決定推薦閾値の高い順に推薦スコアの領域を決定し，これらの領域に対して，それぞれ推薦度を高，中，低と決定することで，以下の通り推薦度を決定することが可能となる．

推薦度高：推薦スコアが8.91以上の番組

推薦度中：推薦スコアが5.38以上8.91未満の番組

推薦度低：推薦スコアが1.63以上5.38未満の番組

本例による推薦番組および推薦度の結果を表4-20に示す．なお，本例では最大推薦アイテム数を設定していないが，実際には最大推薦アイテム数を設定し，推薦アイテム数が最大推薦アイテム数よりも少なくなるように推薦アイテム数決定推薦閾値を設定することも可能である．

表 4-20 推薦計算例による推薦番組および推薦度の結果

Title	Genre	Keywords	Channel	Time	Score	Degree
“Tp”	Drama	“Kc”	1	5	10.78	High
“Tl”	Music	“Kc”	1	4	10.03	
“Tq”	Variety	“Ka”	2	6	7.79	Medium
“Tr”	Drama	“Ka”	4	6	7.52	
“Tn”	Drama	“Ka”	4	4	7.02	
“To”	Drama	“Kb”	3	5	3.74	Low
“Tj”	Variety	“Kb”	2	3	3.51	
“Tk”	Music	“Kb”	3	3	2.99	

実際のテレビにおける実現では、これらの推薦番組及び推薦度を図4-18の通り表示した。図4-18においては、推薦度は星のマークによって表されており、星の数が多いほど推薦度が高いことを示している。すなわち、星の数が3つの場合に推薦度が高いことを示し、2つの場合には推薦度が中程度、1つの場合には推薦度が低いことを示している。

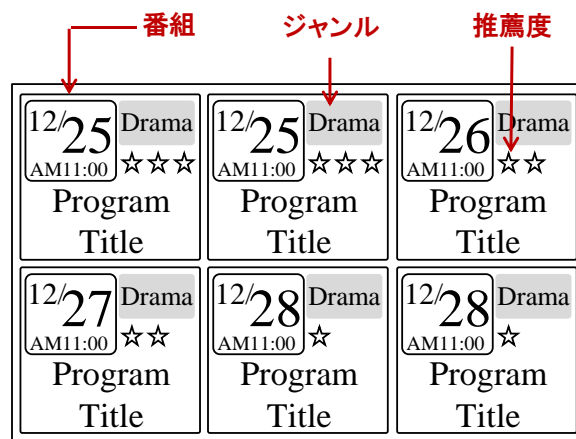


図 4-18 推薦結果および推薦度の表示例

4.5 提案技術の評価と関連技術との比較

実際のテレビ番組のデータを用いて、本提案技術におけるテレビ番組推薦精度と処理負荷を評価した。処理負荷としては300MHzのCPUと256MByteのメモリを用いた場合の処理時間を評価した。また、これらについて関連技術と比較した。以下にこれらの評価結果及び比較結果を示し、考察する。

4.5.1 提案技術の評価

本評価を行うにあたり、1.40GHzで動作するIntel社製のCeleronプロセッサM360及び256MByteのメモリを搭載したWindowsPC(以下、実験PCと呼ぶ)上で本

提案方式によるテレビ番組推薦機能を実現した．評価では24人の一般評価者に2ヶ月間に渡り毎日以下の操作を実施して頂いた．

- (1) 日々のテレビ操作内容（録画，再生，予約，及び検索実績）を記録して頂いた．特に，操作したテレビ番組のタイトルとその操作内容（再生，予約，検索及び録画の種別）を記録及び提出して頂いた．
- (2) 日々操作したテレビ番組のタイトルとその操作内容を実験PC上で実現したテレビ番組推薦機能に入力して頂いた．

評価者の各人には以上の操作を毎日行って頂き，著者らは日々各評価者に上記の処理を実施して頂いた後，各評価者に対して日々以下を実施した．

- (3) (2)で入力頂いたこれまでのテレビ番組の操作内容と翌日からの1週間分のテレビ番組表を基に，実験PC上で実現したテレビ番組推薦機能により1週間分の推薦番組を得た．
- (4) (3)で1週間分の推薦番組得るまでに要した時間（推薦処理時間）を測定して記録した．
- (5) 各評価者に対して，(3)で得られた1週間分の推薦番組の内，翌日分の推薦番組について各評価者にヒヤリングを行った．そして，評価者ごとに推薦に対する正解と不正解及び推薦漏れを判定し，正解数と不正解数及び推薦漏れ数をカウントした．

以上の操作を2ヶ月間に渡り毎日実施し，(5)で得た評価者全員に対する2ヶ月分の正解数(Gn)，不正解数(Bn)，推薦漏れ数(Ln)をそれぞれ合計して，次式により，推薦における適合率(Pr)と再現率(Re)及びF値を計算した．

$$Pr = Gn / (Gn + Bn) \quad (4.15)$$

$$Re = Gn / (Gn + Ln) \quad (4.16)$$

$$F\text{値} = 2 * Pr * Re / (Pr + Re) \quad (4.17)$$

結果，推薦における適合率は82%，再現率は78%，F値は79.9%であった．

また、(4)で測定及び記録した、推薦番組を得るまでに要した時間（推薦処理時間）の中から、最大で要した推薦処理時間を得た．結果、1週間分のテレビ番組の推薦処理に最大で1616ミリ秒を要していた．これは、1.40GHzのCPU及び256MByteのメモリを搭載した実験PCでの結果であるため、300MHzのCPU、256MByteのメモリを持つシステムでは次式により7541ミリ秒を要すると考えられる．

$$1616(\text{ミリ秒}) \times 1400(\text{MHz}) \div 300(\text{MHz}) = 7541 \text{ミリ秒} \quad (4.18)$$

実際に300MHzのCPU、256MByteのメモリを持つ普及型の録画機能付きテレビにて提案技術を実現し、(1)で得られたデータをもとに1番組あたりの処理時間を測定したところ、最大で2.67ミリ秒で処理を完了した．日本における地上波デジタル放送では、1週間に約3000番組のテレビ番組が放送されるため、これらの番組を約8秒で推薦可能となるといえる．

4.5.2 関連技術との比較

比較として、タイトルと内容説明文から抽出したキーワードをカウントし、ユーザが習慣的に視聴している番組に多く含まれるキーワードをユーザの嗜好として学習する方式[61]をベースとして、番組ジャンルとチャンネル[59]及び放送時間帯[60]も考慮したコンテンツベース拡張方式を評価した．本評価を行うにあたり、上述したコンテンツベース拡張方式による推薦機能を、提案方式の評価で用いた実験PC上で試作し、先の(1)から(5)の手順で評価した．特に、(2)において、試作したコンテンツベース拡張方式の推薦機能に、(1)で得られた記録の中から1ヶ月間分、3人の評価者に記録して頂いた日々のテレビ操作内容とテレビ番組表を入力した．そして、(3)において、試作した推薦機能により、日々入力して頂いたテレビ操作内容の翌日から3日分のテレビ番組を推薦した．ここでは3日分のテレビ番組の推薦番組数を10番組に固定し、(5)において翌日の推薦番組に対して、先の評価者3人にヒヤリングを行

った．そして，評価者ごとに推薦に対する正解と不正解及び推薦漏れを判定し，正解数と不正解数及び推薦漏れ数をカウントした．本操作を1ヶ月における毎日分実施し，(5)において先の3人の評価者全員に対する1ヶ月分の正解数(Gn)，不正解数(Bn)，推薦漏れ数(Ln)をそれぞれ合計して，式(4.15)乃至(4.17)により，推薦における適合率(Pr)と再現率(Re)及びF値を計算した．また，(4)で測定及び記録した，推薦番組を得るまでに要した時間(推薦処理時間)の中から，最大で要した推薦処理時間を基に，1番組の推薦計算に要する最大時間を計算した．結果，推薦における適合率は71%であり，再現率は62%であった．また，F値は66.2%であった．さらに，1番組あたりの処理時間は最大で1.85ミリ秒であった．そのため，本方式では，300MHzのCPU，256MByteのメモリにて1番組あたり8.67ミリ秒で処理を完了可能であり，日本における地上波デジタル放送1週間分のテレビ番組(約3000番組)の推薦処理には約26秒必要となるといえる．

また，高速協調フィルタリング方式[84]による推薦機能についても評価した．本評価でも，提案方式の評価で用いた実験PC上で高速協調フィルタリング方式による推薦機能を試作し，先の(1)から(5)の手順で評価した．特に，(2)において，試作した高速協調フィルタリング方式の推薦機能に，(1)で得られた記録の中から評価者24人の1ヶ月間分のテレビ操作内容とテレビ番組表を入力した．そして，(3)において，試作した推薦機能により，評価者24人のテレビ操作内容とテレビ番組表に基づいて，先のコンテンツベース拡張方式の推薦機能の評価で対象とした3人の評価者の推薦番組を得た．具体的には，この3人の評価者に日々入力して頂いたテレビ操作内容の翌日から3日分のテレビ番組を推薦して評価した．ここでも3日分のテレビ番組の推薦番組数を10番組に固定し，(5)において翌日の推薦番組に対して，先の評価者3人にヒヤリングを行った．そして，評価者ごとに推薦に対する正解と不正解及び推薦漏れを判定し，正解数と不正解数及び推薦漏れ数をカウントした．本操作を1ヶ月における毎日分実施し，(5)において先の評価者3人に対する1ヶ月

分の正解数 (Gn), 不正解数 (Bn), 推薦漏れ数 (Ln) をそれぞれ合計して, 式 (4.15) 乃至 (4.17) により, 推薦における適合率 (Pr) と再現率 (Re) 及びF値を計算した. また, (4) で測定及び記録した, 推薦番組を得るまでに要した時間 (推薦処理時間) の中から, 最大で要した推薦処理時間を基に, 1ユーザ及び1番組あたりに推薦計算に要する最大時間を計算した. 結果, 推薦における適合率は60%であり, 再現率は78%であった. また, F値は67.8%であった. さらに, 1ユーザ及び1番組あたり最大で3.21ミリ秒の処理時間を要した. そのため, 本方式では, 300MHzのCPU, 256MByteのメモリにて1ユーザ及び1番組あたり14.98ミリ秒で処理を完了可能であり, 日本における地上波デジタル放送1週間分のテレビ番組 (約3000番組) の推薦処理には約45秒必要となるといえる.

さらに, ハイブリッド方式[78]をベースとして, 高速協調フィルタリング方式[84]と上述のコンテンツベース拡張方式によるハイブリッド推薦機能を評価した. 本評価でも, 提案方式の評価で用いた実験PC上でハイブリッド推薦機能を試作し, 先の(1)から(5)の手順で評価した. 特に, (2)において, 試作したハイブリッド方式の推薦機能に, (1)で得られた記録の中から評価者24人の1ヶ月間分のテレビ操作内容とテレビ番組表を入力した. そして, (3)において, 試作した推薦機能により, 入力した評価者24人のテレビ操作内容とテレビ番組表に基づいて, 先のコンテンツベース拡張方式の推薦機能の評価で対象とした3人の評価者の推薦番組を得た. 具体的には, この3人の評価者に日々入力して頂いたテレビ操作内容の翌日から3日分のテレビ番組を推薦して評価した. ここでも3日分のテレビ番組の推薦番組数を10番組に固定し, (5)において翌日の推薦番組に対して, 先の評価者3人にヒヤリングを行った. そして, 評価者ごとに推薦に対する正解と不正解及び推薦漏れを判定し, 正解数と不正解数及び推薦漏れ数をカウントした. 本操作を1ヶ月における毎日分実施し, (5)において先の評価者3人に対する1ヶ月分の正解数 (Gn), 不正解数 (Bn), 推薦漏れ数 (Ln) をそれぞれ合計して, 式 (4.15) 乃至 (4.17) により,

推薦における適合率 (Pr) と再現率 (Re) 及びF値を計算した。また、(4)で測定及び記録した、推薦番組を得るまでに要した時間 (推薦処理時間) の中から、最大で要した推薦処理時間を基に、1ユーザ及び1番組あたりに推薦計算に要する最大時間を計算した。結果、推薦における適合率は65%であり、再現率は76%であった。また、F値は70.1%であった。さらに、1ユーザ及び1番組あたり最大で5.35ミリ秒の処理時間を要した。そのため、本方式では、300MHzのCPU、256MByteのメモリにて1ユーザ及び1番組あたり24.97ミリ秒で処理を完了可能であり、日本における地上波デジタル放送1週間分のテレビ番組 (約3000番組) の推薦処理には約75秒必要となるといえる。

提案技術についても関連技術と同じ評価者3人、1ヶ月分の記録を用いて評価しなおしたところ、推薦における適合率は78%、再現率は76%、F値は77.0%であった。一方で、処理時間については、先の24人、2ヶ月分での処理時間の方が最大値としては大きくなる可能性があるため、先に示したとおり、1週間のテレビ番組 (約3000番組) の推薦処理には約8秒必要になるものとした。

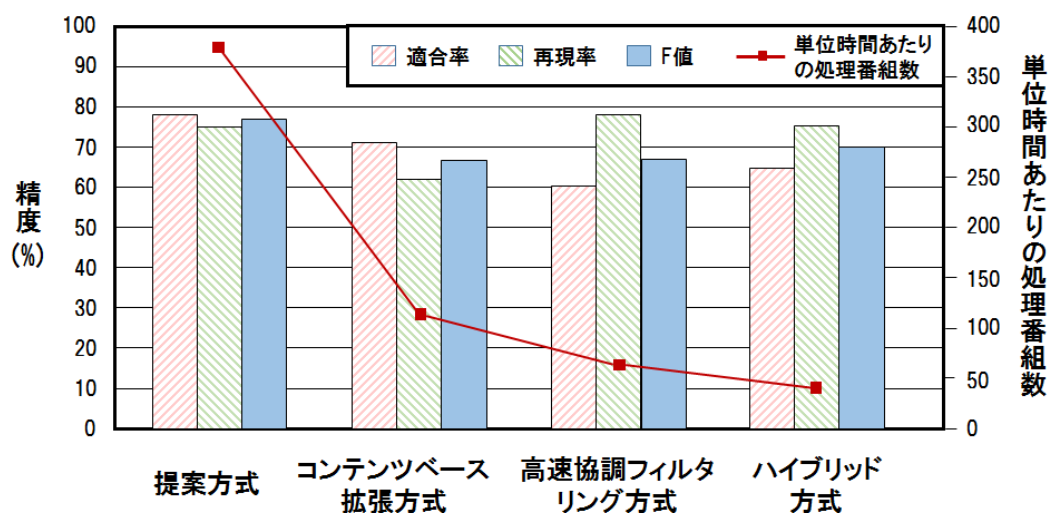


図4-19 関連方式との比較結果

これらの結果を図4-19に示す。図4-19においては、処理負荷として、便宜上300MHzのCPU, 256MByteのメモリを使用した場合の単位時間あたりの処理番組数を示した。

これらの結果からわかるとおり、提案方式はテレビ番組推薦機能における適合率と再現率のバランスを改善すると共に必要計算リソース軽減に向けた処理負荷の低減と高精度化を実現できている。

4.5.3 評価結果及び比較結果に対する考察

関連技術と提案技術における精度と処理時間を第2章で述べた処理負荷と精度の位置づけの図にマッピングすると図4-20のとおりとなる。図4-20では、処理負荷の代わりに、300MHzのCPU, 256MByteのメモリで、日本における地上波デジタル放送1週間分のテレビ番組（約3000番組）に対して推薦計算を行った場合の処理時間を示した。図4-21からわかるとおり、提案技術は他の関連技術では入っていない本研究のターゲットに入っており、必要計算リソースの低減と高精度化を実現できている。

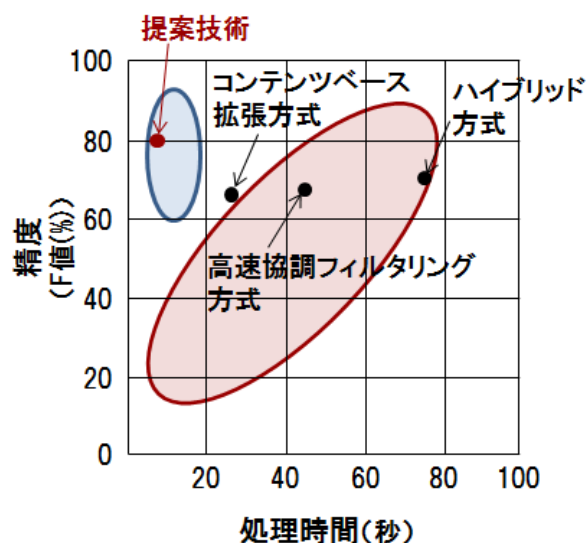


図4-20 処理負荷と精度の位置づけに対するマッピング

また，提案技術において，ある評価者の1日の推薦結果を抽出し，推薦閾値の変化に対する推薦適合率，再現率，及びF値の変化を調べたところ，結果は図4-21のとおりであった．本結果からも，提案手法では推薦における適合率と再現率のバランスをよくすることができていることがわかる．これらにより，関連技術のいずれの方式でも達成することができない図4-1で示した本研究の目標を提案方式により達成することが可能となる．

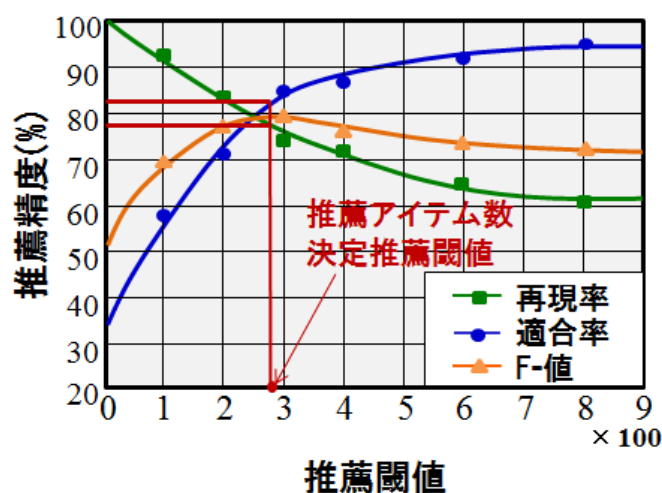


図 4-21 推薦閾値の変化に対する推薦適合率，再現率，及び F 値の変化

4.6 新規性と有効性

本研究の新規性はN-S Curveによる推薦番組数の決定方式と嗜好抽出におけるテレビ番組の属性及び操作の定義にある．また，その有用性は推薦における適合率と再現率のバランスを最適化できる点と推薦に必要となる処理負荷を軽減し，必要計算リソースを低減できる点にある．

従来，推薦方式としては，コンテンツベース方式と協調フィルタリング方式及びこれらを組合わせたハイブリッド方式があった．しかし，これらのいずれの方式においても，推薦番組数に制限を与える場合には，推薦スコアの

高い順にN個に固定しており、推薦における適合率と再現率のバランスをとる推薦番組数としていなかった。このため、これらによるテレビにおける番組推薦機能は、推薦してほしい番組が推薦されない場合や逆に推薦してほしい番組が推薦されてしまうことも多く、ユーザにとっての使い勝手が良くないものとなっていた。また、従来は嗜好抽出に用いる番組属性として、チャンネル、ジャンル、出演者名、キーワード、視聴時間、視聴時間帯、言語、制作国、制作者名、監督名、ディレクター名等、多くの属性が使用されていた。しかし、これらの属性の重みはすべてが同等ではなく、加味することでノイズとなる属性も多い。嗜好抽出に用いるテレビ操作としても、視聴操作と録画操作が用いられていた。しかし、視聴操作の考慮はノイズになりやすく、他にもユーザの嗜好を反映すると考えられるテレビ操作が多かった。推薦処理では、嗜好抽出に用いる番組のキーワード抽出に形態素解析等の多くの処理負荷がかかっていた。このため、必要計算リソースも多くなり、処理負荷及び必要計算リソースの削減は今後の課題とされていた。通常、ユーザがあるリクエストをテレビに対して行った際に、ユーザが待つことができる時間は10秒程度である。このため、推薦処理も10秒以下で完了する必要がある。もちろん、処理を高速に行うには、使用するCPUやメモリを高価なものとするればよいが、これは機能搭載対象とするテレビの原価コストを上げることになってしまう。本研究では、推薦対象のテレビ番組に対するユーザの嗜好度を示す推薦スコアと推薦番組数の関係を示す**N-S Curve**を定義し、推薦番組数を制御することで、テレビ番組推薦における適合率と再現率の最適化を可能とした。加えて、ユーザの嗜好を的確に表現する必要最低限のテレビ操作として、再生、予約、検索、録画を定義した。また、テレビ番組の重要属性として、番組タイトル、番組キーワード、番組ジャンル、放送チャンネル、放送時間を定義し、これらの重み付けによる推薦スコア計算方式を得た。さらに、本方式では番組キーワードとして、従来多くの処理を必要としていた形態素解析を必要としない出演者名等を用いた。これらにより、推薦処理に

おける処理負荷と必要計算リソースを低減すると共に推薦の高精度化を実現した。

本研究成果により、関連技術のいずれの方式でも達成することができない図4-1の目標を達成した。これにより、普及型の録画機能付きテレビにおいて、CPUやメモリを変更及び追加搭載せず、テレビ番組推薦機能を付加可能となる。すなわち、既存の普及型の録画機能付きテレビにおいて、ハードウェア構成を変えず、テレビの原価コストを上げることなく、ユーザにとって使い勝手の良いテレビ番組推薦機能を付加可能となる。

推薦システムでは推薦対象のアイテムに対する推薦スコアを計算するが、これらの推薦スコアが一様に減衰することは稀である。すなわち、これらの推薦システムにおいて、本N-S Curveが線形になることは稀であり、N-S Curveは1つ以上の変曲点を持つと考えられる。このため、N-S Curveの適用は、本研究で実現したテレビ番組の推薦に対してのみならず、一般アイテムの推薦における推薦アイテム数を決定する際にも有効であると考えられる。

4.7 第4章まとめ

本章では、推薦対象のアイテムに対するユーザの嗜好度を示す推薦スコアと推薦アイテム数および推薦精度の関係を明確化し、これを表現する“Number-Score Curve (N-S Curve)”を定義した。また、本N-S Curveを用いて推薦すべきアイテムを決定するための推薦閾値を制御することによって、推薦アイテム数を最適化し、推薦における適合率と再現率のバランスを最適化する方式を提案した。さらに、実際に本N-S Curveを用いたテレビ番組の推薦システムを一般の評価者に使用して頂くことで評価し、その有効性を確認した。

本研究の新規性はN-S Curveによる推薦番組数の決定方式と嗜好抽出におけるテレビ番組の属性及び操作の定義にある。また、その有用性は推薦における適合率と再現率のバランスを最適化できる点と推薦に必要となる処理負

荷を軽減し、必要計算リソースを低減できる点にある。本成果により、普及型の録画機能付きテレビにおいて、CPUやメモリを変更及び追加搭載せず、ハードウェア構成を変えることなく、すなわちテレビの原価コストを上げることなく、テレビ番組推薦機能を付加可能とした。

現状、推薦システムでは推薦対象のアイテムに対する推薦スコアを計算しており、これらの推薦スコアが一様に減衰することは稀である。すなわち、これらの推薦システムにおいて、N-S Curveが線形になることは稀であり、N-S Curveは1つ以上の変曲点を持つと考えられる。このため、N-S Curveの適用は、本研究で実現したテレビ番組の推薦に対してのみならず、一般アイテムの推薦における推薦アイテム数を決定する際にも有効であり、テレビ番組推薦機能のみならず、商品や情報などの一般的なアイテムを推薦する推薦システムの実現にも有効であると考えられる。

今後は、アイテムの属性および操作に対する定義および重み値の設定方法を最適化することで、さらなる推薦精度の向上を図る。また、N-S Curveにおける変曲点をより高速かつ最適に決定できる方式を検討することで、より効果的に推薦閾値を決定し、推薦精度を向上可能な方式へと発展させていく。

第5章

今後の展望

本研究では，ユーザの好みに応じたテレビ番組の推薦技術及び当該テレビ番組内のシーン検索技術を提案した．しかし，ユーザの好みは世の中の流行を生み出すトレンドにも影響されると考えられる．そこで，本章では，今後の展望として，本研究で提案したユーザの好みに基づく推薦/検索機能に対して，世の中のトレンドを加味した推薦/検索へと拡張するためのトレンド把握方式について検討する．特に，近年普及が目覚ましいソーシャルメディアに着目し，消費者のマインドセット（思考や価値観等）が多分に含まれているものとして活用が期待されているソーシャルメディアデータの活用によるトレンド把握方式について検討する．具体的には，「いつ」，「何を」，「どのように」推薦/検索するかを明確にするために，ユーザのアイテム選択履歴データとソーシャルメディアデータを融合して世の中のトレンドを把握する．これにより，流行る可能性のあるアイテムの特定し，トレンド変動要因の把握した上で，「いつ」，「何が」，「どのくらい」流行るかを予測するための需要予測方式を検討する．

5.1 はじめに

前章までは，コストに対する要求の厳しい普及型の録画機能付きテレビにおいて，シーン検索機能及びテレビ番組推薦機能を付加可能とする技術を提

案してきた。一方で、近年、ネットワーク接続機能を持ったテレビも出てきている。現状ではまだネットワークへの接続率は低いものの、今後この接続が一般化した場合には、サーバにこれらの機能を持たせることで、低コストで多様な機能をテレビに提供可能となる。

第3章及び第4章で提案した技術は高精度かつ低処理負荷でシーン検索機能及びテレビ番組推薦機能を実現できる。このため、提案した技術はサーバに適用された場合でもシーン検索機能及びテレビ番組推薦機能における使い勝手の良いユーザインタフェースや高速処理を実現可能であり、有効な技術である。機能的にも、これらの機能により、ユーザの好みに応じたテレビ番組の推薦及び当該テレビ番組内のシーンの検索を行うことができるため、ユーザにとって有効である。しかしながら、ユーザの好みは変わっていくものであり、これらの好みは世の中の流行を生み出すトレンドにも影響されると考えられる。このため、推薦や検索においても、ユーザの好みだけでなく、世の中のトレンドをいち早く把握して、このトレンドに基づいて推薦や検索を行うことが有効であると考えられる。

世の中のトレンドは多くの人々のつながりや行動に依存しており、トレンドを把握するには自分以外の人々の行動も分析する必要がある。テレビがネットワークに繋がっているならば、サーバにおいて多くの人々のつながりや行動を基にトレンドを把握できる可能性がある。特に、近年のネットワーク上ではソーシャルメディアが普及しており、ソーシャルメディアのデータを活用することで、世の中のトレンドを把握できる可能性がある。

次節以降では、今後の展望として、前章までに提案したユーザの好みに基づく推薦/検索機能に対して、サーバにおけるソーシャルメディアデータの活用による世の中のトレンドを加味した推薦/検索への拡張を検討する。まず、ソーシャルメディア及びそれに関連した環境の現状を俯瞰し、広くソーシャルメディアデータ利活用の可能性を述べる。次に、ソーシャルメディアデータの活用による世の中のトレンドの把握可能性について検討する。

5.2 ソーシャルメディア環境の現状

近年，インターネットを介して，人との繋がりをサポートするメディアとして，ソーシャルメディアが普及しており，誰もがこのメディアを介して自由に情報を発信できるようになった．Twitter[86]やFacebook[87]等に代表されるSNS(Social Networking Service)は，ソーシャルメディアを実現するネットワークサービスであり，数億人以上のユーザを保有している．これらのユーザは友人や家族とのコミュニケーションや，同じ趣味嗜好を持つユーザとの情報交換等にSNSを活用している．

このような背景から，これらのソーシャルメディア或いはSNSで受発信される情報（以降，ソーシャルメディアデータ或いはSNSデータと呼ぶ）には，利用者の実体験に基づく，商品やサービスに対する感想や評価が数多く含まれている．そのため，ソーシャルメディアデータには，消費者のマインドセット（思考や価値観等）が多分に含まれており，このデータの活用に対する期待が高まっている．特に，ソーシャルメディアデータは消費者の行動や世の中のトレンドを把握するための有力な情報源の一つとして，その活用が期待されている．

本節では，ソーシャルメディアデータの活用による世の中のトレンドの把握可能性を検討するにあたり，まずは現状のソーシャルメディア関連環境を整理する．

5.2.1 ソーシャルメディアの利用状況

近年，SNSの利用者は急速に増加しており，例えば，TwitterのMAU(Monthly Active Users)は2013年12月の時点で全世界で約2億4100万人に達している[88]．また，FacebookのMAUは，2014年4月の時点で約12億8000万人に達し

ている[89]. さらに, 2011年6月にサービスを開始したLINEの躍進も著しく, 2014年4月の時点で世界での利用者が4億人に達しており[90], 今後も利用者が増加していくものと考えられる. このようなことから, SNSの注目度は高まっており, 消費者の声を理解するデータソースとしても, その活用が期待されている.

5.2.2 ソーシャルメディアの利用目的

SNSの利用者は, 現状, 目的別にSNSを使い分けている. 例えば, 家族や友人と連絡をとるためのツールとして, 或いは, 調べた情報や調べたい情報等の興味のある情報を受発信ためのツールとして, 利用目的にあわせて, それぞれ使いやすいSNSを選択して利用している[91]. 例えば, 家族や友達との連絡手段として, LINEやFacebook等を利用しているケースが多く見られる. また, 調べたい情報や興味のある情報を得るためにTwitter等を利用しているケースが多く見られる. このように, SNSはコミュニケーション及び情報受発信ツールとして利用されている.

5.2.3 ユーザの利用機器と行動の変化

利用者が使用するIT(Information Technology)機器も変化しており, 情報通信端末の主役が従来のPC(Personal Computer)やテレビ等から携帯電話からスマートフォンやタブレット型の端末へと変化してきている[91]. このようなスマートフォンやタブレット端末の普及は, 今後も進むと考えられ, いつでもどこでも大量の情報のやり取りが可能になり, ユーザの行動にも影響してくるものと考えられる. テレビ番組等の動画も今後はスマートフォンやタブレット端末等のモバイル端末でも多く視聴されるようになると予想される.

5.2.4 市場への影響

マーケティング手法にも変化がみられ、従来のマスメディアによる広告から、インターネットとスマートフォン向けの広告宣伝活動の比重が増加してきている[91]。特に、SNSの広告メディアとしての市場規模は、現状では1,000億円程度であるが、今後数年で現在の雑誌広告並みの規模へと成長する可能性がある。これらの傾向から、今後、インターネット広告の中でも特にSNSがメディアとしての存在価値を増していくと考えられる。マーケティングの効果としても、SNSの位置づけは拡大してきている。例えば、消費者に対するアンケート結果によると、商品やサービスの購入の際に参考とする情報源の上位に友人や家族のクチコミやSNSが位置づけられている[92]。このような背景から、ソーシャルメディアの活用により消費者の意向を汲み取る活動が活発化してきている[93][94]。

5.2.5 ソーシャルメディア活用のメリットとデメリット

前述の通り、ソーシャルメディアでは、人との繋がりをサポートしており、利用者も多く、誰もが自由に情報を発信できるため、クチコミ効果が絶大である。消費者はソーシャルメディアを利用して自由に希望や不満等を発言することができるため、ソーシャルメディアデータを分析することで、消費者のニーズや意見を獲得することができる可能性がある。さらに、ソーシャルメディア上での人々の日常的な会話から、これまで知りえなかった消費者のインサイトを発見できる可能性がある。

一方で、ソーシャルメディアの活用にはデメリットも潜在する。例えば、現状では、発言内容のチェックが十分に行われていないことも多い。このため、ソーシャルメディアデータにはデマや不確実な情報が含まれていることも多く、これを利用することで、誤ったニーズやインサイトを得てしまう可

能性もある。したがって、ソーシャルメディアデータの分析には十分な注意が必要となり、これらのリスク対策として、デマや不確実な情報の発見技術が積極的に研究されている[95]。

5.3 ソーシャルメディアデータ利活用の可能性

前節で述べたとおり、近年、ソーシャルメディアの利用数は増加しており、ソーシャルメディアデータには消費者のマインドセットが含まれている可能性がある。このため、ソーシャルメディアデータを活用することで、世の中のトレンドを把握できる可能性がある。そこで、本研究では、ソーシャルメディアデータの活用により、世の中のトレンドを把握する方式を検討する。そして、世の中のトレンドに基づくテレビ番組や当該テレビ番組におけるシーンをはじめとする情報やアイテム（以下、単にアイテムと呼ぶ）の推薦/検索の可能性を検討する。

これまで、アイテムの推薦や検索を、現時点でのユーザの好みに基づいて行ってきた。特に、推薦機能の実現では、ユーザのアイテムに対する選択履歴を用いてきた。アイテムに対する選択履歴を分析することで、ユーザの好みを把握し、好みに合ったアイテムを推薦できる。しかしながら、これらの履歴データ利活用では、ユーザの好みに対する現状を理解できても将来を予測することは困難であるといった課題がある。一方で、ユーザの好みは変わっていくものであり、これらの好みは世の中のトレンドに影響されると考えられる。これらの課題に対して、ソーシャルメディアデータをアイテムの選択履歴データと組み合わせて利活用することで、世の中のトレンドに合ったアイテムをいち早く推薦あるいは検索できる可能性がある。

現状、ソーシャルメディアデータを推薦や検索に活用する技術がいくつか提案されている。具体的には、ユーザが興味・関心のある情報や検索エンジンで検索する際のキーワードをソーシャルメディアデータから入手する技術

が提案されている[96]. 本技術では, SNSにおける人のつながりを利用しており, 相手に何か情報を伝える際には, 相手がほしがる情報を伝えていることに立脚し, 口コミで伝搬している新たな情報を推薦することが可能となる. また, 同様に情報の伝搬に着目した技術として, SNSにおけるコミュニティでの情報の伝搬を利用した推薦技術[97]や製品のレビューサイトにおける信頼性のあるつながり及び情報に基づく推薦システムも提案されている[98]. さらに, ソーシャルメディアデータを分析することで, ユーザの隠れた興味に基づく意外性のある推薦を実現する技術も提案されている[99]. テレビ番組の推薦に関しても, SNSでは, 現在見ている番組に対するコメントも書き込まれることに着目して, SNSからユーザどうしの関係や他ユーザの書込みを取得し, 協調フィルタリングにより番組推薦する手法が提案されている[100]. これにより, テレビ番組の視聴履歴データを取得できないユーザも対象とすることが可能となる. 世の中のトレンドに着目した推薦技術としても, CGM(Consumer Generated Media)での発信情報に基づいて世の中のトレンドを分析し, このトレンドに基づいた情報推薦技術が提案されている[101]. しかしながら, 本技術では, 推薦対象とするアイテムが「いつ」「どのくらい」流行するかについては考慮していない. また, トレンド語に基づく職業における業種の推薦であり, 流行している或いは流行しそうなアイテムそのものを推薦する技術ではない. したがって, 推薦対象とするアイテムに関するデータとの融合分析を行うものではなく, 「どのような時」「何を推薦すべきか」については考慮されていない.

本節では, ソーシャルメディアデータ活用した世の中のトレンドに基づくアイテムの推薦と検索の可能性を検討する. これを実現するために, ソーシャルメディアデータと推薦対象のアイテムに関する選択履歴データとの融合分析により, 世の中のトレンドを把握する方式を検討する. 特に, 世の中のトレンドを把握するにあたり, 流行る可能性のあるアイテムの特定し, トレンド変動要因の把握する. そして, 「いつ」「何を」「どのように」推薦あ

るいは検索可能とするかを明確にするために，世の中のトレンドを反映したアイテムの需要予測を行うことで，「いつ」「何が」「どのくらい」流行るかを予測する．

5.4 トrend把握方式の検討

本節では，ソーシャルメディアデータとアイテムの選択履歴データの融合分析により，世の中のトレンドを把握する方式を検討する．以下では，トレンドを把握するにあたり，検討対象のデータ，流行る可能性のあるアイテムの特定，トレンド変動要因の把握，及び，トレンドを反映した需要予測について述べる．

5.4.1 検討対象データ

本研究の対象はテレビ番組であるため，本来であれば，テレビ番組のデータとソーシャルメディアデータとの融合分析により，テレビ番組やテレビ番組内のシーンの検索に向けた世の中のトレンド把握可能性を検討したいところである．しかし，テレビ番組に関する情報の活用は著作権の観点からもその活用が限られている．このため，今回はアイテムの選択履歴データとして，飲料のPOS(Pint of Sales)データを使用し，世の中のトレンドを売上で置き換えて検討した．また，ソーシャルメディアデータとして，実際のTwitterデータを使用した．以下にこれらのデータの出典とデータ期間を示す．

《POSデータ》

- ・日経メディアマーケティング社から購入した全国約800店舗の飲料（約1300品目）の売上データ
- ・データ期間：2011年2月～2013年1月

《ソーシャルメディアデータ》

- ・ Twitter社の Streaming APIで収集した実ツイートデータ（全ツイートデータの1%）
- ・ データ期間：2012年7月～2013年1月

5.4.2 流行る可能性のあるアイテムの特定

前述の通り，ソーシャルメディアでは，商品やサービスに対する感想や評価が数多く発言されている．このため，ソーシャルメディアにおける発言数が商品の売上に少なからず影響を及ぼしていると考えられる．しかしながら，

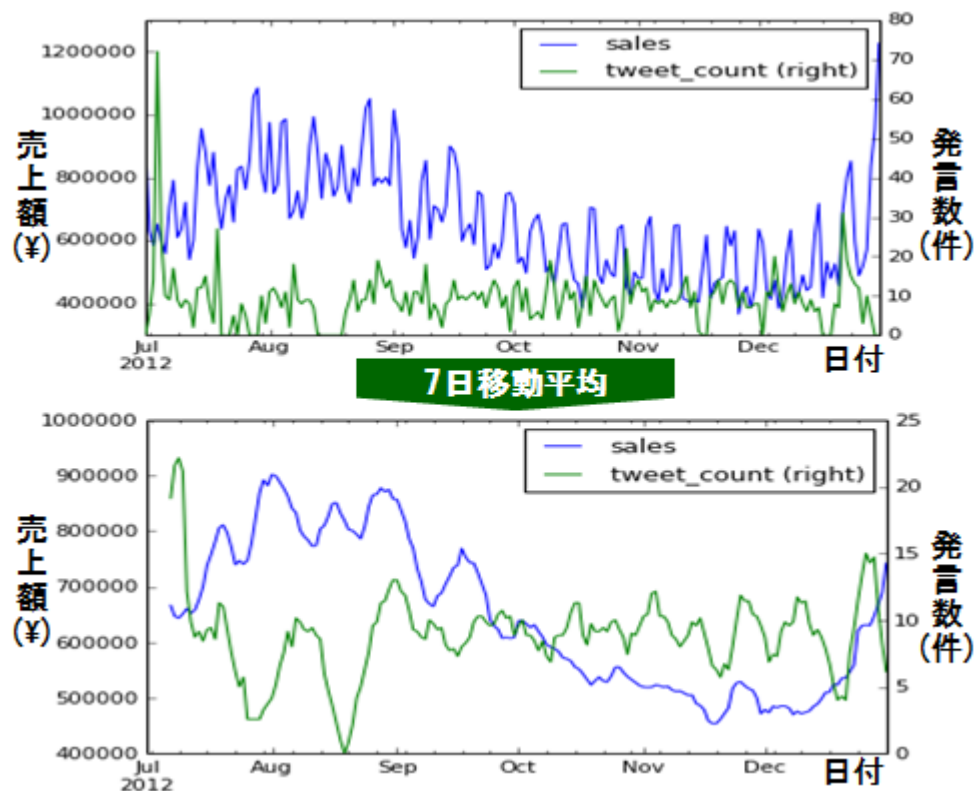


図5-1 定番アイテムにおける売上額とTwitter発言数の時系列推移

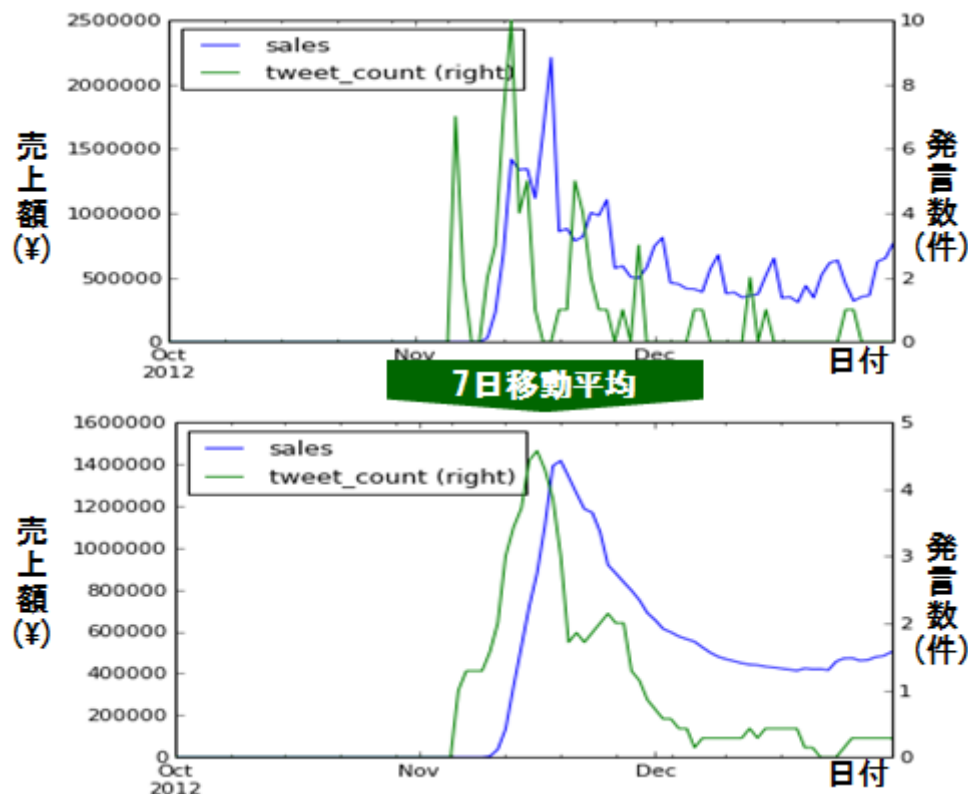


図5-2 新商品における売上額とTwitter発言数の時系列推移

実際には、消費者によく知られた定番商品と、ユーザにとって目新しい新商品で、ソーシャルメディアにおける発言数と商品の売上の相関性は異なっている。図5-1及び図5-2は、それぞれ、ある定番商品（飲料A）及び新商品（飲料B）における、POSデータ上の日別の売上額とTwitter発言数の時系列推移である。図5-1からも分かる通り、飲料Aの売上額は周期性の変動を持ち、Twitter上の発言数との相関性は低いことが分かる。

一方、飲料Bの売上額は、発売後、数日から2週間程度でピークとなり、その後は緩やかな下降傾向を示す。また、Twitterでは発売前から話題となり、発売直後にも飲料Bに対する感想等を含む発言が増加する傾向があり、飲料Aと比較して、売上額とTwitter発言数の間の相関性が高い。特に、Twitter発言数を5日遅らせた時系列推移との相関係数（ $R=0.913$ ）が最大となった。

ただし、全ての新商品において、売上とTwitter発言数の間に相関があるわけではない。図5-3は、実際に2012年7月～12月に新発売された94商品について、商品発売前のTwitter発言数と売上の関係を示した図である。

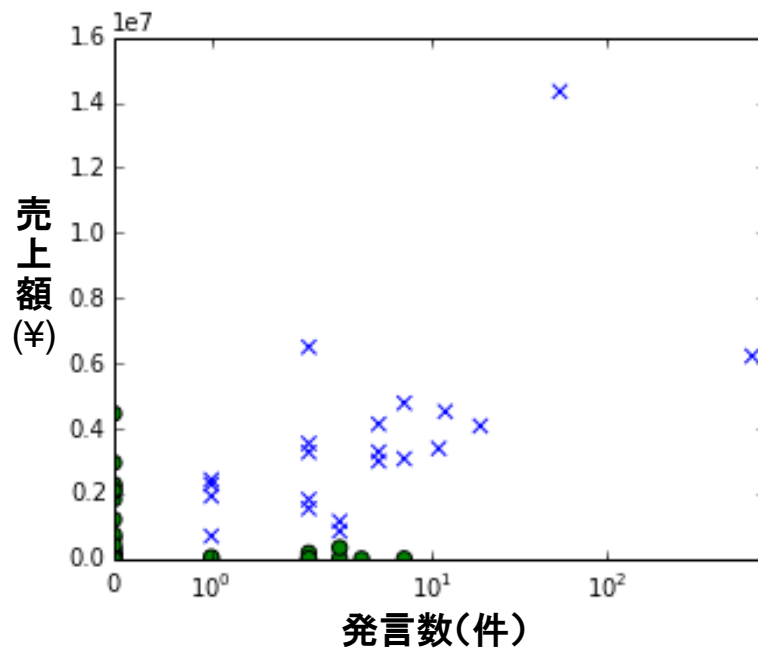


図5-3 商品発売前のTwitter発言数と売上の関係

図5-3に示す通り、×印で示した商品（21商品）は売上額とTwitter発言数に関連性が見られる。一方、●印で示した商品（73商品）はTwitter発言数に対して売上額が少ない、或いはTwitter発言数が0件であるにも関わらず売上額が多く、関連性が見られない。これは、前者のケースでは、商品名が「ラムネ」や「サイダー」等の単純なものであり、発売会社の商品とは関係のない発言がカウントされたためであると考えられる。一方、後者のケースでは、商品名が長いため、Twitter上では発言されにくい商品、或いは今回使用したTwitter社提供のStreaming APIで収集できるデータが、Twitterでの発言における全体の1%であったために、関連発言データを取得できなかった可能性がある。しかしながら、少なくとも発売前にTwitter発言数の増大が観測された場

合には、売上額も増大する可能性があり、新商品に関する売上変動要因をTwitter発言数から捕捉できる可能性があると考えられる。

5.4.3 トレンド変動要因の把握

前節で新商品に関する売上額とTwitter発言数の間に相関性が見られたため、以降では新商品について考察を進める。商品の売上は、Twitter発言数だけでなく、評判にも深い関連があると考えられる。そこで、先の新商品（飲料B）について、売上額とTwitter上での評判（ポジティブな意見とネガティブな意見）の関係を調べた。その結果を図5-4に示す。

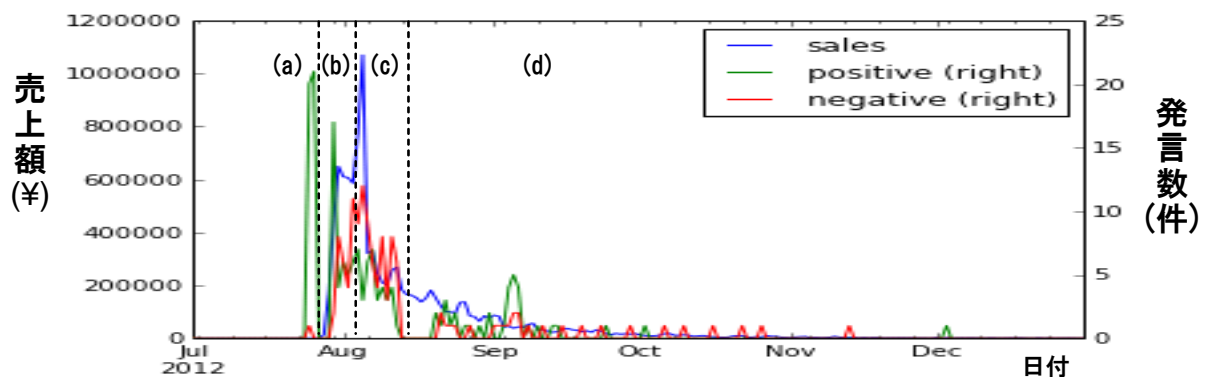


図5-4 新商品についての売上とTwitter上での評判の関係

図5-4において、新商品発売前（(a)の期間）では、この商品に対して、「飲んでみたい」「おいしそう」等のポジティブな意見が増加している。発売直後（(b)の期間）でも、依然として「意外といける」「クセになる」等のポジティブな意見が増加しており、これらの意見と共に、売上も増加している。しかしながら、売上ピーク時を過ぎたころ（(c)の期間）から、「まずい」「おすすめしない」等のネガティブな意見が増加し、それと共に売上も減少し始めている様子が分かる。さらに、衰退期（(d)の期間）では、「いつまで飲め

るだろうか」等の中立的意見が度々見られる程度で、発言数は少なくなり、売上も下降している。このように、アイテムに関するSNS上のデータの感情を分析或いは本文を確認することで、トレンド変動要因を把握できると考えられ、これらを捕捉して、推薦等の施策を講じることができると考えられる。

5.4.4 トレンドを反映した需要予測

以上の結果を基に、トレンドを反映した需要予測方式を検討した。特に、新商品に対して、過去の類似商品の販売実績を基にしたフェルミ推定[102]にSNS上の発売前の発言数を加味して売上規模を予測した。この結果を図5-5に示す。ここで、SNS上の発売前の発言数を加味するに当たり、Twitter発言数に関する回帰式を生成し、需要予測対象の新商品における発売前のTwitter発言数と生成した回帰式を用いて新商品の売上額を予測した。

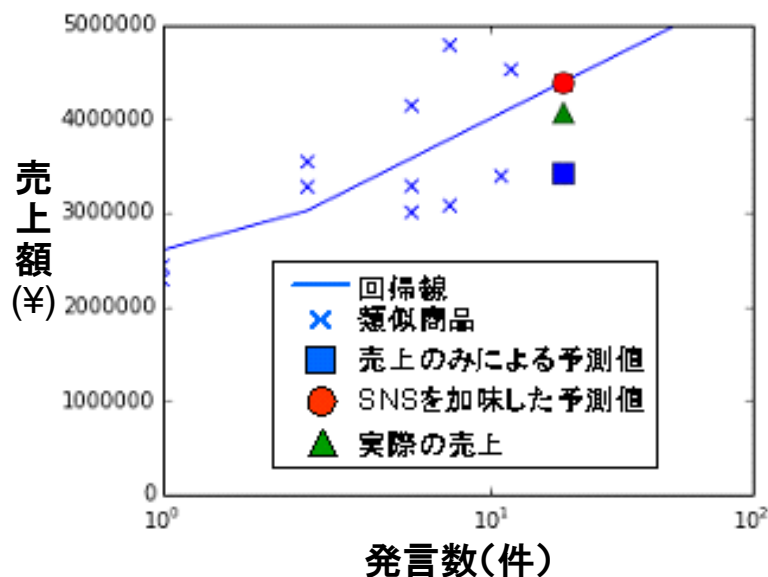


図5-5 新商品の需要予測結果例

この結果、従来の売上のみからの予測手法（類似商品の平均）では、実際の売上との誤差が15.9%であったのに対し、Twitterでの発言数を加味した手

法では、当該誤差を7.6%とすることができた。

また、本手法による売上推移の予測結果を図5-6に示す。図5-6からも分かる通り、売上推移についても、ピーク時に誤差があるものの、上昇/下降トレンドは一致しており、妥当な需要予測結果を得ることができた。

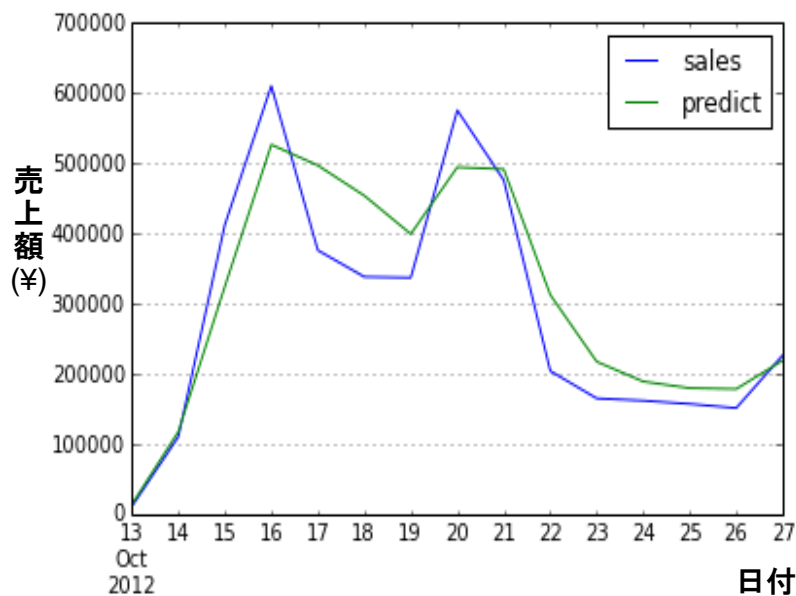


図5-6 売上推移の予測結果

5.5 トレンド予測結果の評価

図5-3の×印で示した商品（21商品）に対して、従来の売上のみからの予測手法とTwitterでの発言数を加味した手法とで売上予測を行った。それぞれについて、発売2週間後の予測値と実際の売上額との誤差を図5-7に示す。図5-7は、それぞれの手法での予測値の誤差が実際の売上額に占める割合を商品毎に算出した結果である。図5-7において、斜線塗りで示しているグラフが従来の売上のみからの予測手法による結果であり、色塗で示したグラフがTwitterでの発言数を加味した手法による結果である。

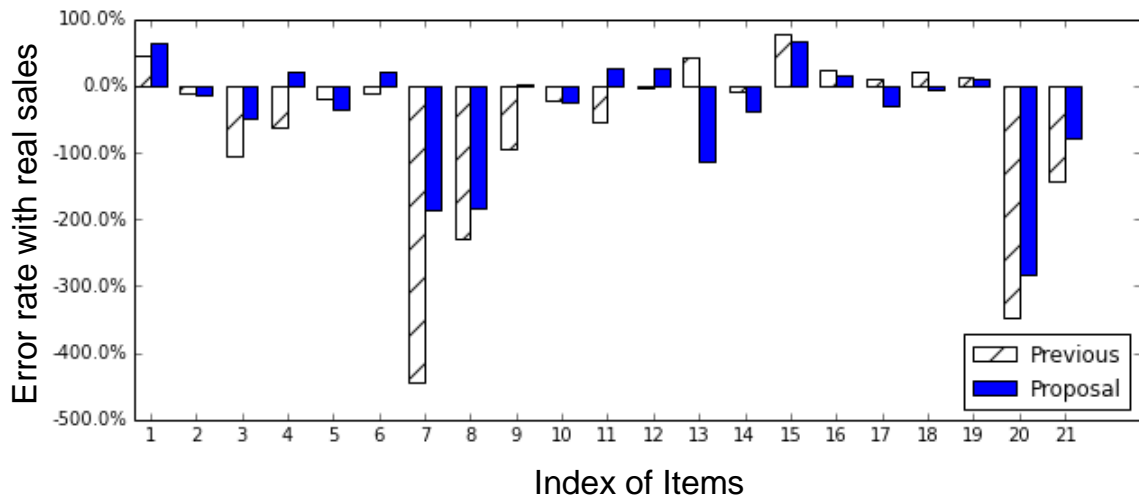


図5-7 需要予測誤差の評価

図5-7から分かる通り，全21商品のうち16商品で精度改善が見られた．また，21商品の誤差の平均は，従来の売上のみからの予測手法で63.1%，Twitterでの発言数を加味した手法で37.7%となり，Twitterでの発言数を加味することで，誤差の平均を25.4%低減することができた．また誤差の分散を63.8%低減することができた．これらの結果から，ソーシャルメディアデータと推薦対象のアイテムに関する選択履歴データとの融合分析により，世の中のトレンドを把握することができると考えられる．特に，流行る可能性のあるアイテムの特定，トレンド変動要因の把握，及びトレンドを反映したアイテムの需要予測を行うことで，「いつ」「何が」「どのくらい」流行るかを予測することができると考えられる．これにより，「いつ」「何を」「どのように」推薦あるいは検索可能とするかを判断可能になると考えられる．

5.6 第5章のまとめ

ソーシャルメディアデータには，利用者の実体験に基づく，商品やサービ

スに対する感想や評価が数多く含まれている．そのため，ソーシャルメディアデータには，消費者のマインドセット（思考や価値観等）が多分に含まれており，このデータの活用に対する期待が高まっている．本研究では，ソーシャルメディアデータを活用したトレンドに基づくアイテムの推薦/検索の可能性について検討した．特に，ユーザのアイテム選択履歴データとソーシャルメディアデータの融合分析により，世の中のトレンドを把握する方式として，流行る可能性のあるアイテムの特定し，トレンド変動要因の把握した上で，アイテムの需要予測方式を検討した．

ソーシャルメディアデータの活用には，サーバへのネットワーク接続が不可欠となる．一方で，現状ではまだ接続率は高くないものの，ネットワーク接続機能を持ったテレビも市場に出てきている．そのため，今後テレビにおいてネットワーク接続が一般化した場合に，前章までで提案した技術をサーバで活用すると共に，ソーシャルメディアデータの活用も可能となる．これにより，テレビに対して低コストで使い勝手が良く，高速処理が可能なテレビ番組推薦機能及びシーン検索機能を提供可能となる．また，これまでのようにユーザの好みに基づく推薦や検索だけでなく，世の中のトレンドにいち早く対応して推薦や検索を行う機能に拡張することが可能となる．

ソーシャルメディアデータには，デマや不確実な情報も多分に含まれている可能性がある．そのため，今後はデマ情報の利用によるトレンド把握及び推薦への影響を検討し，これらによる判断の誤りに対するリスク対策も検討していく．また，ソーシャルメディアでは，発言されにくい商品も多く，逆に関係の無い情報（ノイズ）も多い．そのため，今後は適切なキーワードの設定及びノイズ除去方法も検討していく．そして，ユーザの好みや世の中のトレンドだけでなく，ソーシャルメディアデータを用いた推薦対象者の社会的つながり（ソーシャルグラフ）及び興味上でのつながり（インタレストグラフ）も加味した推薦を実現していく（図5-8）．

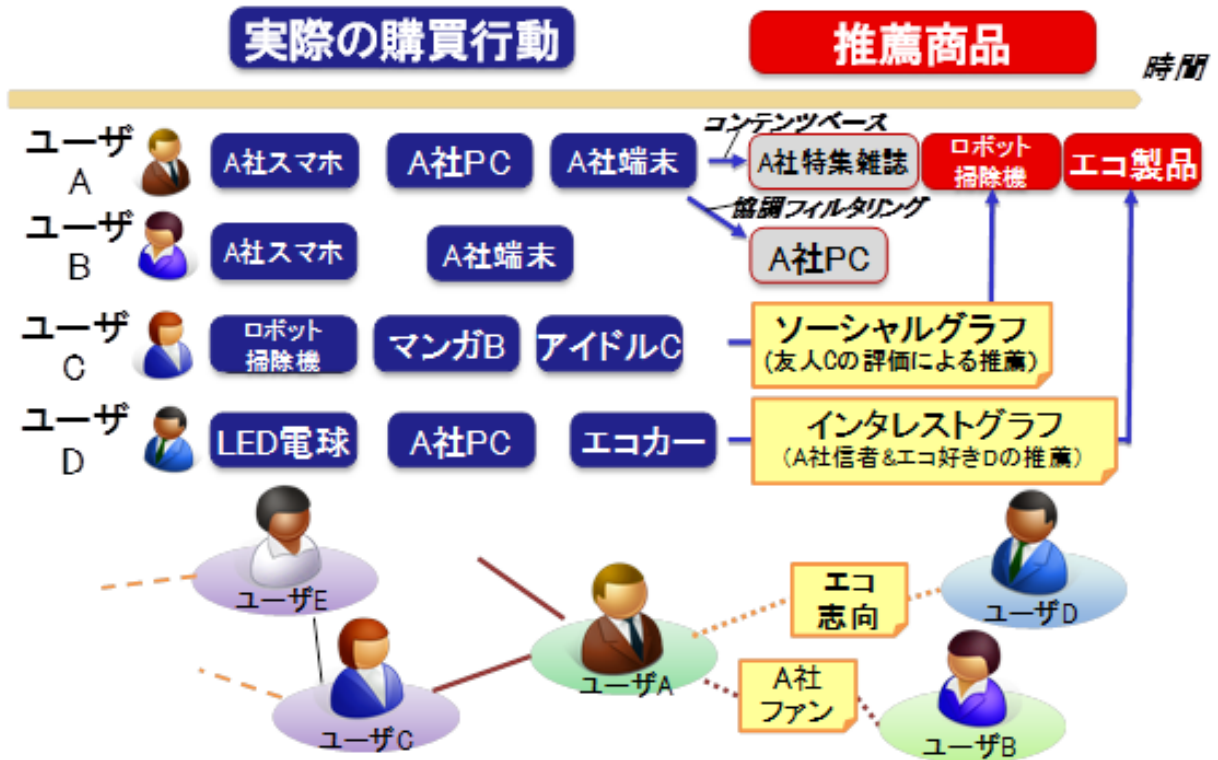


図5-8 ソーシャルグラフ及びインタレストグラフによる推薦

第6章

まとめ

本研究では，テレビの新視聴スタイルを実現する技術として，近年大量化するテレビ番組を見落とし無く，短時間で視聴可能とする技術を提案した．特に，ユーザが好みのシーンを選択して視聴する「シーン検索機能」とユーザの好みに合ったテレビ番組を推薦及び自動録画する「テレビ番組推薦機能」の実現技術を開発した．本研究成果により，コストや信頼性及び使い勝手に対する要求が厳しく，処理能力やユーザインターフェースに制限がある普及型の録画機能付きテレビの原価コストを上げることなく，「シーン検索機能」と「テレビ番組推薦機能」を付加可能とした．本章では，本研究に対する結言と今後の課題を述べる．

6.1 結言

テレビは一般家庭に普及して以来，子供からお年寄りまで，誰もが簡単に楽しめる家庭の中心的存在である．その歴史的背景や身近さから，テレビの設計におけるコストや信頼性及び使い勝手に対する要求仕様は，PC (Personal Computer) 等の他のデジタルメディア機器に比べて非常に厳しい．ユーザインターフェースはリモコンが主体であり，ボタンによる操作が前提となっている．リモコンで離れた場所からテレビを操作することも多く，視覚上の表示解像度が低い．CPUやメモリも可能な限り低価格のものとしており，

余剰能力を持たない．このため、テレビはコストを最適化した設計になっており、最大処理負荷時の処理能力に余裕がなく、他の処理はほとんど行えない設計となっている．一方で、今後も新しいテレビを市場に投入し、より良いコンセプトを消費者に提供するために新たな付加機能を実現する必要がある．しかし、テレビに付加機能を実現する場合には、価格を抑えるために設計変更を少なくし、CPUやメモリを変えないことも多い．また、ユーザに長く安心して使用してもらえる機能とするために、使い勝手や精度の良い信頼性のある機能とすると共にリモコンで簡単に扱える機能とする必要がある．

従来、シーン検索機能では、キーワードや顔画像が指定された場合に、そのキーワードや顔が出現するシーンのみを検出して視聴可能としていた．このため、指定されたシーンに関するトピックの開始位置からズレた位置から再生され、ユーザにとっては精度が良い検索とは言えなかった．テレビ番組推薦機能では、推薦番組数を多くすると、ユーザは多くの選択肢を得ることが可能となるが、好みではない番組も推薦される可能性があり、所望の番組を選択することが困難となる．一方で、推薦番組数が少ない場合には、ユーザは好みの番組を選択することが容易となるが、選択肢が少なく、所望の番組が推薦されない可能性がある．すなわち、これらはトレードオフであり、視覚上の解像度が低い画面に表示された推薦番組をリモコンで簡単に選択可能とするには、最適な推薦番組数でこれらのバランスをとることが重要となるが、従来のテレビ番組推薦機能ではこれらのバランスが考慮されていなかった．

そのため、本研究では、以下の主眼点と問題意識に基づいてシーン検索技術とテレビ番組推薦技術を実現した．

<主眼点>

- (1) 普及型の録画機能付きのテレビにおいて、CPUやメモリを変更或いは追加することなく、シーン検索機能及びテレビ番組推薦機能を付加可能

とする。

- (2) シーン検索において、ユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とする。また、テレビ番組の推薦において最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とする。

<問題意識>

上記の主眼点で述べた項目を実現しないとコストや信頼性及び使い勝手に対する要求仕様が厳しく、処理能力やユーザインターフェースに制限があるテレビにおいて、目的とするシーン検索機能とテレビ番組推薦機能を付加することが困難となる。特に、主眼点(1)の項目については、普及型のテレビではCPU及びメモリは全体の1%ほどしか使えず、これ以下でシーン検索機能を実現できるようにする必要がある。また、テレビ番組推薦機能では、普及型のテレビに備わっている少ないCPU及びメモリ資源で推薦処理を高速に実施し、リクエストに対してユーザが待つことができると考えられる10秒以下で推薦結果を提示できるようにする必要がある。これらを実現しなければ、CPUやメモリを高速なものに変更或いは追加する必要がある、テレビの原価コストが上がってしまう。このため、普及型のテレビにシーン検索機能及びテレビ番組推薦機能を付加することが許容されない可能性がある。また、主眼点(2)の項目については、信頼性がなく使い勝手の良くない機能であると判断され、ユーザに使用されない機能となる可能性があり、シーン検索機能及びテレビ番組推薦機能をテレビに付加するという判断にはなりえない可能性がある。

これらの問題点を解決し、主眼点を実現するため、以下のとおりシーン検索技術とテレビ番組推薦技術を提案した。

6.1.1 シーン検索技術

テレビ番組に付随する字幕情報を利用することで、低コストにシーンキーワードを抽出し、シーンのインデクシングを行う方式を提案した。特に、テレビ番組のジャンル毎に共通のキーワードからなる定型句辞書と、番組に固有のキーワードからなる動画固有辞書による多重インデクシング方式を提案した。これら二種類の意味レベルの異なる辞書により、字幕情報利用で発生し得る検索位置ズレを補正し、ユーザが指定したシーンに関するトピックの最初からズレることなく正確に視聴可能とした。また、キーワードの符号化による計算リソース削減方式を提案した。本研究成果の新規性は字幕情報を用いた多重インデクシング方式とシーンキーワードの符号化にある。また、その有用性は字幕情報利用で発生し得る検索位置ズレを補正できる点と、インデクシングにおける処理負荷を軽減し、必要な計算リソースを低減できる点にある。

従来技術では、シーンのインデクシング処理に8MHzのCPU負荷と2MByteのメモリ量を必要とし、シーン検索精度を示すF値は65.8%程度である。提案方式により、シーン検索精度を示すF値を81.5%とすることができた。また、インデクシングに必要となるCPU負荷を1.835MHz, 必要メモリ量を0.384MByteとすることができた。これにより、普及型の録画機能付きテレビにおいて、CPUやメモリを変更或いは追加することなく、シーン検索機能を付加可能とした。

6.1.2 テレビ番組推薦技術

ユーザの嗜好を的確に表現する必要最低限のテレビ操作として、再生、予約、検索、録画を定義した。また、テレビ番組の重視属性として、番組タイトル、番組キーワード、番組ジャンル、放送チャンネル、放送時間を定義した。そして、これらの重み付けによりテレビ番組に対するユーザの嗜好度を

示す推薦スコアの計算方式を提案した．さらに，推薦スコアと推薦アイテム数の関係を示す“Number-Score Curve (N-S Curve)”により推薦番組数を制御することで，テレビ番組推薦における適合率と再現率の最適化手法を提案した．これにより，最適な推薦番組数でユーザが好む番組を過不足なく推薦可能とした．本研究の新規性はN-S Curveによる推薦番組数の決定方式とユーザの嗜好抽出における番組の重要属性及びテレビ操作の定義にある．また，その有用性は推薦における適合率と再現率のバランスを最適化できる点と，推薦における処理負荷を軽減し，必要となる計算リソースを低減できる点にある．従来技術では，300MHzのCPU，256MByteのメモリを搭載した普及型の録画機能付きテレビにおいて，一週間の番組(3000番組)に対する推薦処理に26秒を要する．また，推薦適合率と再現率は，それぞれ71%及び62%(F値=66.2%)である．これに対して提案方式では，同様の条件下で推薦処理を8秒で実施可能とした．また，テレビ番組推薦における適合率と再現率をそれぞれ82%及び78%(F値=79.9%)とすることができた．これによりテレビ番組推薦機能における適合率と再現率のバランスを改善し，普及型の録画機能付きテレビにおいて，CPUやメモリを変更或いは追加することなく，テレビ番組推薦機能を付加可能とした．

本研究では，さらに，近年様々な分野で活用が検討されているソーシャルメディアデータを活用した推薦の可能性について検討した．特に，ユーザのアイテム選択履歴データとソーシャルメディアデータの融合分析による，世の中のトレンドの把握可能性を示した．これにより，今後テレビのネットワーク接続が一般化した場合に，本研究で提案したテレビ番組推薦機能及びシーン検索機能サーバを適用すると共に，トレンドを加味した推薦/検索を行うことで，低コストで使い勝手が良く，高速処理が可能なテレビ番組推薦機能及びシーン検索機能を提供可能となる．また，これまでのようなユーザの嗜好に基づくだけでなく，世の中のトレンドにいち早く対応して推薦/検索を行うことが可能となる．

以上のとおり，本研究では，消費者にとって身近である一方で処理能力とユーザインターフェースに制限がある普及型のテレビにおいて，好きなテレビ番組を見落としなく素早く選択して，そのシーンを短時間で視聴可能とした．これにより，近年多忙化する生活者が低価格なテレビを用いて，短時間で大量のテレビ番組を楽しむことができる．

新たな付加機能の必要性はテレビに限らずHDR（Hard Disk Recoder）やSTB（Set-Top-Box）でも同じである．また，コンシューマ製品に限らず，サーバ等のプロバイダ製品でも同じである．一方で，本研究の成果で実現できるシーン検索及びテレビ番組推薦における高精度化や必要計算リソース軽減に向けた処理負荷の低減は，対象機器におけるユーザインタフェースの向上や信頼性及び高速処理につながる．特に，本研究では，これらの機能を機能実現要件の厳しい普及型のテレビにおいて付加可能とした．このため，本研究で得られた成果はテレビに限らず多くの製品に対して有効である考えられる．

6.2 今後の展開

本研究では，シーン検索機能とテレビ番組推薦機能をそれぞれ別個の機能として実現可能としたが，今後はこれらを連動した機能として実現していく．すなわち，検索したシーンのキーワードに基づいてテレビ番組の推薦を可能とする．あるいは逆に，推薦したテレビ番組のキーワードに基づいてシーン検索における提示キーワードを生成可能とする．これらにより，シーン検索及びテレビ番組推薦における更なる精度や機能及び使い勝手の向上をめざす．

また，上述したとおり，本研究成果はHDR（Hard Disk Recorder）やSTB（Set-Top-Box）でも活用可能であり，今後はこれらの機器にも適用していく．さらに，本研究成果はコンシューマ製品に限らず，サーバ等のプロバイダ製品にも有効であるため，これらの装置にも適用を推進していく．特に，近年ではまだ接続率は高くないものの，ネットワークへの接続が可能なテレビも

市場に出てきており，サーバを活用してこれらの機能を実現することで，より低コストにこれらの機能をテレビに提供可能とする．

サーバでは，本研究で対象としたテレビとは異なり，計算リソースが充実している可能性がある．先に述べたとおり，提案技術はこのようなサーバで実現された場合でも，実現機能におけるユーザインタフェースや信頼性の向上及び高速処理につながるため有効である．一方で，使用計算リソースを問わず，最高の精度を提供する技術も有効になると考えられる．そのため，今後は，本研究でターゲットとした必要計算リソースの制限を考慮せずに，最高の精度を提供できる技術についても検討していく（図6-1）．そして，計算リソースに制限がある場合には，その制限に合わせた技術の選択により，対象とするシステムに合った技術を提供できるように技術のラインナップを拡充していく．

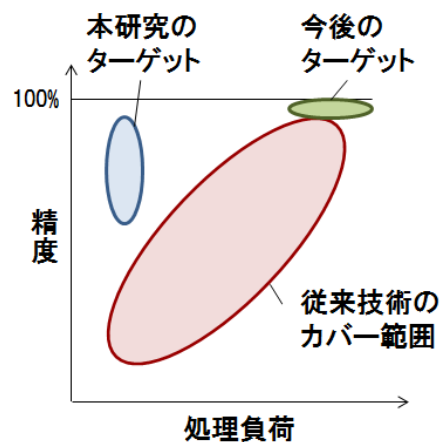


図6-1 処理負荷と精度の位置づけにおける今後のターゲット

また，本研究ではユーザの好みに基づく推薦/検索機能を実現したが，今後はサーバにおいてソーシャルメディアデータを活用することで，世の中のトレンドも反映した推薦/検索機能を実現していく．さらに，推薦機能としては，これまで，図6-2に示す通り，コンテンツベース推薦及び協調フィルタリング推薦による好みの「モノ」や「ヒト」に基づく推薦を行っていた．今後は，

世の中のトレンドや推薦対象者の社会的なつながり，及びユーザの興味上でのつながりといった「コト」による推薦も可能とする「未来志向型」の推薦機能を実現していく．



図 6-2 未来志向型の推薦

謝辞

本研究を進めるにあたっては、多くの方々に大変御世話になりました。研究を順調に進められたことは一重にご協力頂いた皆さまのおかげと考えており、ここに感謝の辞を述べさせていただきます。

本研究を直接御指導頂いた森田啓義教授には終始熱心に御指導を賜りました。論文のまとめ方や表現方法で悩んでいるときには的確な御指導頂けただけでなく、研究課題や進め方まで御指導頂けたことに深く感謝申し上げます。

御忙しい中、査読を快く引き受けてくださいました加藤聰彦教授、吉永努教授、田野俊一教授、栗原聡教授に心より御礼申し上げます。また、貴重なご意見を頂いた電気通信大学大学院情報システム研究科の笠井裕之准教授並びに学生諸氏に感謝いたします。

社会人（特にマネージャ）である私にとって、会社生活と学生生活を両立することは、簡単ではありませんでしたが、学会発表や論文執筆に追われている際に、職務上のマネジメント業務を快くサポートして下さった株式会社日立製作所横浜研究所の藤林昭氏、石黒正雄氏、前岡淳氏、及び株式会社日立ハイテクノロジーズの禰寝義人氏と岡山祐孝氏に深く感謝申し上げます。

会社での多岐にわたる研究及びプロジェクトをリードする立場でもあった私のリーダ業務をご支援頂いた株式会社日立製作所横浜研究所の親松昌幸氏と福井大輔氏に厚く御礼を申し上げます。

電気通信大学という素晴らしい研究環境をご紹介下さった株式会社日立製作所横浜研究所主管研究員の山崎眞見氏、本学への入学を認めて下さった同社同研究所長の山足公也氏ならびに前所長の堀田多加志氏、組込みシ

謝辞

ステムセンタ長の伊藤浩道氏，総務課諸氏に感謝申し上げます．

本研究関連発表を快く御承諾くださり御助言並びに御支援くださった株式会社日立製作所の勝又賢治氏，岡村巧氏，株式会社HGSTジャパンの藤井由紀夫氏，株式会社リコーの川口敦生氏，株式会社日立システムズの田代卓氏，本研究に関する業務をお手伝い頂いた田中晶氏，関本信博氏，株式会社HGSTジャパンの鈴木誠人氏，多忙中，いつも私を励ましてくださった数多くの友人諸氏に感謝いたします．

最後に，博士号取得という大きな挑戦を見守りつつ天国へ旅立った母と，そのような中でも最後まで支援を続けてくださった父並びに家族に，心より感謝いたします．

関連論文の印刷公表の方法及び時期

- (1) 全著者名：廣井和重，森田啓義

論文題目：サブタイトル情報に基づくシーンの多重インデクシング
方式と高効率テレビ視聴システムの提案

印刷公表の方法及び時期：映像情報メディア学会誌68巻8号，
pp. J323-J328（2014年8月）

（第3章に関連）

- (2) 全著者名：Kazushige Hiroi and Hiroyoshi Morita

論文題目：High-accuracy time-variant recommender system with
Number-Score curve

印刷公表の方法及び時期：Proceedings of 2014 IEEE Fourth
International Conference on Consumer Electronics -
Berlin (ICCE-Berlin), pp.106-110, Sep.7-10, 2014

（第4章に関連）

参考論文の印刷公表の方法及び時期

- (1) 全著者名：廣井和重，田代卓，川口敦生

論文題目：VLIW型メディアプロセッサを用いたMPEG-2オールフォーマットビデオデコーダ

印刷公表の方法及び時期：映像情報メディア学会誌56巻5号，
pp. 804-813（2002年5月）

- (2) 全著者名：廣井和重

論文題目：ソーシャルメディアデータ利活用の可能性

印刷公表の方法及び時期：情報知識学会誌23巻4号，pp. 462-472（2013
年12月）

その他の発表及び講演

(1) 技術報告

廣井和重，田代卓，川口敦生，“VLIW型メディアプロセッサを用いたMPEG-2ビデオデコーダ”，映像情報メディア学会技術報告，Vol. 25, No. 21, pp. 25-30, 2001

(2) 年次大会

廣井和重，森田啓義，“サブタイトル情報を用いた低処理負荷シーンマルチインデクシング方式と高効率テレビ視聴システムの開発”，映像情報メディア学会年次大会講演予稿集，“2-2-1”-“2-2-2”，2013

(3) 招待講演

廣井和重，“ソーシャルメディアデータ利活用の可能性”，情報処理学会 連続セミナー2013東京（同時遠隔放送：大阪）

関連主筆出願特許

【名称】	【出願日】	【公開番号】	【登録番号】
(1) 動画処理装置	2005/4/19	P2006-303746	P04525437
(2) 動画処理装置	2005/4/19	P2010-171999	
(3) 動画処理装置	2006/3/7	US2006/0233522	
(4) 動画処理装置	2006/3/20	CN1856065	PZL200610065519.3
(5) 動画処理装置	2006/3/20	CN101959043	
(6) 動画再生装置	2006/1/13	P2007-189473	
(7) 動画再生装置	2006/11/15	US2007/0168867	
(8) 動画再生装置	2006/12/11	CN101001345	
(9) 動画再生装置及び動画再生方法			
	2006/8/23	P2008-053877	P04730253
(10) 動画再生装置	2006/12/12	P2008-148077	P04905103
(11) 動画再生装置	2004/12/24	P2006-180305	P04349277
(12) 動画再生装置	2004/12/24	P2009-095055	
(13) 動画再生装置	2004/12/24	P2011-109675	
(14) 動画再生装置	2005/6/17	US2006/0140580	P7796857
(15) 動画再生装置	2005/6/17	US2010/0329636	
(16) 動画再生装置	2005/7/18	CN1794798	PZL200510085425.8
(17) 動画再生装置	2005/7/18	CN101600074	PZL200910136925.8
(18) 動画記録再生装置			
	2004/12/24	P2006-180306	

- (19) 動画記録再生装置
2004/12/24 P2009-055637 P04760893
- (20) 動画記録再生装置
2004/12/24 P2011-101393
- (21) 動画記録再生装置
2005/12/22 US2006/0147184 P8218945
- (22) 動画記録再生装置
2005/12/26 CN1832557
- (23) 動画記録再生装置
2005/12/26 CN101835018 PZL201010193465.5
- (24) 動画記録再生装置
2005/12/26 CN102611863
- (25) 動画再生装置 2007/6/7 US2008/0138034
- (26) 動画再生装置 2007/12/12 CN101202864 PZL200710194201.X
- (27) C M検出方法及びこれを用いた動画再生装置
2007/11/7 P2009-118204 P04929127
- (28) 動画処理装置及び動画処理方法
2007/11/7 P2009-118205
- (29) 動画処理装置及び動画処理方法
2008/11/6 US2009/0116816
- (30) 動画再生装置 2007/11/7 P2009-118207 P04871839
- (31) 動画インデクシング方法及び動画再生装置
2009/4/7 P2010-245853
- (32) 動画インデクシング方法及び動画再生装置
2010/3/23 US2010/0257156
- (33) 動画インデクシング方法及び動画再生装置
2010/4/6 CN101859586

- (34) 動画再生装置 2010/3/30 P2011-211481
- (35) 動画再生装置 2011/1/14 CN102208205
- (36) 動画推薦システム及び動画推薦方法
2010/6/24 P2012-008789
- (37) 動画推薦システム及び動画推薦方法
2011/3/3 US2011/0320471
- (38) 動画推薦システム及び動画推薦方法
2011/3/4 CN102300126
- (39) 動画推薦装置及び動画推薦方法
2011/3/28 P2012-203773
- (40) 動画推薦装置及び動画推薦方法
2012/1/10 US2012/0254906
- (41) 動画推薦装置及び動画推薦方法
2012/1/13 CN102710973
- (42) 情報検索システム及び情報検索方法
2011/12/27 P2013-134755

参考文献

- [1] 内閣府, 消費動向調査 平成26年11月調査結果, 2014.
- [2] 鈴木宏幸, 中嶋満雄, 浜田宏一, 都留康隆, 及川勇太, “デジタルテレビ「Wooo」の高画質, エコへの取り組み,” 日立評論, Vol.93, No.10, pp.57-61, 2011.
- [3] 丸山隆, 安達聡, 四本直樹, 廣井和重, 岩村真澄, 清水喜弘, 武井健, “次世代Priusを支える技術,” 日立評論, Vol.87, No.10, pp.35-38, 2005.
- [4] 林昭夫, 廣井和重, 水野達吏, 永松健司, “新しいインフラストラクチャーに対応したブロードバンドパソコン “Prius” ,” Vol.87, No.10, pp.31-34, 2005.
- [5] 長坂晃朗, 田中譲, “カラービデオ映像における自動索引付け法と物体探索法,” 情報処理学会誌, 33-4, pp. 543-550, 1992.
- [6] 親松昌幸, 廣井和重, 森靖英, 加藤雅弘, 木村淳一, “自動ダイジェスト再生機能のためのリプレイシーン検出手法,” 情報科学技術フォーラム一般講演論文集, Vol.6, No.3, pp.189-190, 2007.
- [7] L.Pearlstein, J.Henderson and J.Boyce, “An SDTV Decoder with HDTV Capability : ”An All-Format ATV Decoder”,” Proceedings of 137th SMPTE Technical Conference and World Media Expo, pp422-434, 1995.
- [8] J.Boyce and L.Pearlstein, “Low-cost All Format ATV Decoding with Improved Quality,” 30th SMPTE Advanced Motion Imaging Conference, pp.45-51, 1996.
- [9] 小島啓二, 西岡清和, 野尻徹, 宮崎健司, “メディアプロセッサ(MAP)のビジョンとアーキテクチャ,” 情報処理学会全国大会講演論文集, 第59回平成11年後期(特別1), pp.87-92, 1999.

- [10] 鈴木教洋, 川口敦生, 田代卓, “メディアプロセッサ(MAP)とメディアライブラリ,” 情報処理学会全国大会講演論文集, 第59回平成11年後期(特別1), pp.93-94, 1999.
- [11] 宮崎太郎, 中川弘充, 中川竜太, 岩野公司, 篠田浩一, 古井貞熙, “野球中継番組を対象とした音響情報を用いたシーン認識,” 日本音響学会2006年春季講演論文集, pp.19-20, 2006.
- [12] 関口一樹, 小杉信, 向井信彦, “映像と音情報を用いた野球中継の自動インデクシング,” 電子情報通信学会技術研究報告, 画像工学, Vol.106, No.398, pp.41-46, 2006.
- [13] 安藤亮一, 篠田 浩一, 古井貞熙, 望月貴裕, “動画像インデクシングのためのシーン時系列の確率的言語モデル,” 第12回画像センシングシンポジウム予稿集, pp.513-518, 2006.
- [14] Jiang Feihu, Hiroyoshi Morita and Akiko Manada, “Semantic Analysis of Structured High-definition MPEG-2 Soccer Video Using Bayesian Network,” Information Systems International Conference, pp.483-490, 2013.
- [15] 佐藤真一他, 電子情報通信学会『知識の森』2群(画像・音・言語)－11編(マルチメディア)「2章 マルチメディアコンテンツ解析」, 電子情報通信学会, 2012.
- [16] 佐藤真一他, 電子情報通信学会『知識の森』2群(画像・音・言語)－11編(マルチメディア)「4章 マルチメディアアプリケーション」, 電子情報通信学会, 2012.
- [17] Rainer Lienhart, “Automatic text recognition for video indexing,” Proceedings of the fourth ACM international conference on Multimedia, pp.11-20, 1996.
- [18] Toshio Sato, Takeo Kanade, Ellen K. Hughes and Michael A. Smith., “Video OCR for digital news archive,” Proceedings of International Workshop on Content-Based Access of Image and Video Database, pp.52-60, 1998.

- [19] Jae-Chang Shim, Chitra Dorai and Ruud Bolle, “Automatic text extraction from video for content-based annotation and retrieval,” Proceedings of International Conference on Pattern Recognition, pp.618-620, 1998.
- [20] Rainer Lienhart and Frank Stuber, “Automatic text recognition in digital videos,” Proceedings of SPIE 2666, Image and Video Processing IV, pp.180-188, 1996.
- [21] 平松義崇, 関本信博, 森靖英, 新庄広, 丸川勝美, “テロップ認識のための領域配置情報を用いた文字領域抽出,” 電子情報通信学会総合大会講演論文集, 情報・システム(2), P.201, 2008.
- [22] Nilesh Bhojne, Pravinkumar Kamde and S.P.Algur, “News Video Indexing and Retrieval using Overlay Text,” Computer Science & Information Technology, pp.11-17, 2012.
- [23] Trung Quy Phan, Palaiahnakote Shivakumara, Shangxuan Tian and Chew Lim Tan, “Recognizing Text with Perspective Distortion in Natural Scenes,” ICCV2013, pp.569-576, 2013.
- [24] 平井辰典, 中野倫靖, 後藤真孝, 森島繁生, “シーンの連続性と顔類似度に基づく動画コンテンツ中の同一人物登場シーンの同定,” 映像情報メディア学会誌, Vol.66, No.7, pp.J251-J259, 2012.
- [25] 野宮浩揮, 宝珍輝尚, “簡潔な顔特微量の選択と統合による映像からの効率的な表情表出シーン検出,” 電子情報通信学会技術研究報告 Vol.111, No.76, pp. 37-42, 2011.
- [26] 平井辰典, 中野倫靖, 後藤真孝, 森島繁生, “音楽動画コンテンツ中のアーティスト名とその登場シーンの同定方法,” 情処研報音楽情報科学, 2012-MUS-94-24, pp.1-8, 2012.
- [27] 久保田英俊, 桃崎浩平, 青木恒, 風間久, “映像シーンを簡単に検索できる” 顔deナビ” ,” 東芝レビュー, Vol.63, No.11, pp.54-57, 2008.
- [28] 内海ゆづ子, 松本雄大, 岩井儀雄, “分岐限定法を用いた顔認証の効率

- 化,” 情報処理学会論文誌, Vol.51, No.2, pp.217-228, 2010.
- [29] 内海ゆづ子, 坂野悠司, 前川敬介, 岩村雅一, 黄瀬浩一, “局所特微量と投票処理を用いた大規模データベースに対する高速顔認識,” 電子情報通信学会論文誌, Vol.J97-D, No.8, pp.1263-1272, 2014.
- [30] 大附克年, 別所克人, 水野理, 松尾義博, 松永昭一, 林良彦, “音声認識を用いたマルチメディアコンテンツのインデクシング,” 情報処理学会研究報告, 2003-SLP-47, pp.19-24, 2003.
- [31] Daniel Povey, Arnab Ghoshal, Gilles Boulianne, Lukas Burget, Ondrej Glembek, Nagendra Goel, Mirko Hannemann, Petr Motlicek, Yanmin Qian, Petr Schwarz, Jan Silovsky, Georg Stemmer, Karel Vesely, “The Kaldi Speech Recognition Toolkit,” Conference paper of IEEE 2011 Workshop on Automatic Speech Recognition and Understanding, pp.1-4, 2011.
- [32] 菅原隆伸, 何光霽, 藤永剛史, 宮本優貴, 野口紘希, 和泉慎太郎, 川口博, 吉本雅彦, “6万語彙実時間連続音声認識のための40nm,144mW音声認識専用プロセッサの開発,” 電子情報通信学会技術研究報告, pp.79-84, 2011.
- [33] 宮本優貴, 何光霽, 和泉慎太郎, 川口博, 吉本雅彦, “2.4倍速実時間6万語彙連続音声認識プロセッサの開発,” 電子情報通信学会技術研究報告, Vol.112, No.365, pp.49-53, 2012.
- [34] 何光霽, 宮本優貴, 松田薫平, 和泉慎太郎, 川口博, 吉本雅彦, “3×倍速実時間6万語彙連続音声認識のための40-nm, 54-mW音声認識専用プロセッサ,” 電子情報通信学会技術研究報告, Vol.113, No.237, pp.29-34, 2013.
- [35] 河合吉彦, 住吉英樹, 八木伸行, “電子番組表における紹介文を利用した番組紹介映像の自動生成手法,” 電子情報通信学会論文誌, Vol.J91-D, No.8, pp.2157-2165, 2008.
- [36] 山田一郎, 佐野雅規, 住吉英樹, 柴田正啓, 八木伸行, “アナウンサーと
-

- 解説者のコメントを利用したサッカー番組セグメントメタデータ自動生成,” 電子情報通信学会論文誌, Vol.J89-D, No. 10, pp.2328-2337, 2006.
- [37] 三浦菊佳, 山田一郎, 住吉英樹, 八木伸行, 奥村学, 徳永健伸, “放送番組を素材としたマルチメディア百科事典の自動構築,” 映像情報メディア学会誌, Vol.62, No.1, pp.110-116, 2008.
- [38] 三浦菊佳, 山田一郎, 小早川健, 松井淳, 後藤淳, 住吉英樹, 柴田正啓, “番組分割に向けたクローズドキャプション中の反復句抽出,” 電子情報通信学会技術研究報告, NLC, 言語理解とコミュニケーション, pp.53-58, 2009.
- [39] 山下道生, “番組シーン再生のための字幕情報を用いた検索技術,” 東芝レビュー, Vol.69, No.4, pp.50-53, 2014.
- [40] 奥村学, *自然言語処理の基礎*, コロナ社, 2010.
- [41] 三上崇志, 相川勇之, 川又武典, “ダブル配列を用いた文書検索用キーワード提示の高速化,” 情報処理学会研究報告, 2009(35), pp.61-67, 2009.
- [42] 谷治隆史, 関口一樹, 向井信彦, 小杉信, “テロップ情報を用いた野球中継ダイジェスト映像の自動生成,” 電子情報通信学会技術研究報告, Vol.105, No.431, pp.29-34, 2005.
- [43] 井手一郎, 浜田玲子, 田中英彦, 坂井修一, “字幕の語義に基づくニュース映像分類,” 第57回情報処理学会全国大会, No.4V-4, Vol.3, pp.259-260, 1998.
- [44] Ichiro Ide, Koji Yamamoto and Hidehiko Tanaka, “Automatic Video Indexing Based on Shot Classification,” AMCP'98, LNCS 1554, pp.87-102, 1999.
- [45] Shin'ichi Satoh, Yuichi Nakamura and Takeo Kanade, “Name-It: Naming and Detecting Faces in News Videos,” IEEE Multimedia, 6, No.1, pp.22-35, 1999.
- [46] Keita Yamamuro and Katunobu Itou, “Speaker Labeling Using Closed-Captioning,” IARIA, pp.38-42, 2011.
- [47] 山室慶太, 伊藤克亘, “デジタル放送の字幕情報と発話傾向を考慮した
-

- 発話者アノテーション,” 情報処理学会全国大会講演論文集, 2012(1), pp.619-621, 2012.
- [48] 谷口行信, 南憲一, 佐藤隆, 桑野秀豪, 児島治彦, 外村佳伸, ”SceneCabinet: 映像解析技術を統合した映像インデクシングシステム,” 電子情報通信学会論文誌, Vol.J84-DII, No.6, pp.1112-1121, 2001.
- [49] 桑野秀豪, 松尾義博, 川添雄彦, “映像・音声認識, 自然言語処理の適用によるメタデータ生成の作業コスト削減効果に関する考察,” 映像情報メディア学会誌, Vol.61, No.6, pp. 842-852, 2007.
- [50] 立林和夫, 入門タグチメソッド, 科技連出版社, 2004.
- [51] 土方嘉徳, “嗜好抽出と情報推薦技術,” 情報処理, Vol.48, No.9, pp.957-965, 2007.
- [52] Dietmar Jannach, Markus Zanker and Gerhard Friedrich, *Recommender systems*, Tutorial at the 23rd International Joint Conference on Artificial Intelligence Beijing, 2013.
- [53] 折原良平, 村上知子, 坂本典哉, 堀口健生, “コンテンツの高精度推薦技術によるデジタル機器の価値向上,” 東芝レビュー, Vol.61, No.12, pp.13-17, 2006.
- [54] 小野智弘, 麻生英樹, 本村陽一, “情報・コンテンツのレコメンド技術と課題,” 電子情報通信学会誌, Vol.94, No.4, pp.310-315, 2011.
- [55] 神 嶌 敏 弘 , 推 薦 シ ス テ ム の ア ル ゴ リ ズ ム , <http://www.kamishima.net/archive/recsysdoc.pdf>, 2014.
- [56] 官原浩二, 小谷亮, 小川吉大, 小林啓二, 近藤省造, “利用者の視聴履歴に基づくTV 番組推薦システムの検討,” 情報処理学会第54回全国大会, 4-245, pp.245-246, 1997.
- [57] 土屋誠司, 佐竹純二, 近間正樹, 上田博唯, 大倉計美, 蚊野浩, 安田昌司, “TV番組推薦システムの構築とその有用性の検証,” 情報処理学会研究会報告, HI-117(14), pp.95-102, 2006.
-

- [58] 高間康史, 難波広樹, 岩瀬徳宏, 服部俊一, 武藤優樹, 庄司俊寛, "テレビ視聴時の情報推薦に基づくヒューマン・ロボットコミュニケーション," 人工知能学会全国大会 (第21回) 論文集, 2D5-5, pp.1-2, 2007.
- [59] 黒木修隆, 廣瀬裕二, 鈴木達也, 片岡充照, 沼昌宏, 山本啓輔, "テレビ視聴者の選局行動に基づく番組嗜好度の推定," 映像情報メディア学会誌, Vol.60, No.3, pp.454-457, 2006.
- [60] 井川一樹, 福原知宏, 藤井秀樹, 武田英明, "テレビ番組の視聴履歴と電子番組表を用いた番組推薦システムの構築と評価," 人工知能学会全国大会 (第24回) 論文集, 3C4-3, pp.1-4, 2010.
- [61] 小川修太, 西本喜則, 大盛善啓, "テレビ番組推薦の精度を向上させる番組メタデータ補完技術," 東芝レビュー, Vol.69, No.7, pp.52-55, 2014.
- [62] 上江まり子, 橋本隆子, 北川博之, "再生リストを利用したWeb 動画クラスタリング手法の提案," DEIM Forum 2012, F6-1, 2012.
- [63] YouTube : <http://www.youtube.com/>
- [64] 吉尾透, 太田学, "ユーザの繋がりを用いた意外性のある動画推薦システム," DEIM Forum 2012, B4-1, 2012.
- [65] 北村祐太郎, 澤勢一史, 延原肇, "形式概念分析を用いた推薦理由を明示する動画推薦手法," 日本知能情報ファジィ学会誌, Vol.25, No.1, pp.624-635, 2013.
- [66] Barry Smyth and Paul Cotter, "Surfing the Digital Wave: Generating Personalized TV Listings Using Collaborative," Case-Based Recommendation, Proceedings of the International Conference on Case-Based Reasoning Berlin, Germany: Springer-Verlag, pp.561-571, 1999.
- [67] Barry Smyth and Paul Cotter, "Personalized Electronic Program Guide for Digital TV," AI Magazine, Vol.22, No.2, pp.89-98, 2001.
- [68] Kamal Ali and Wijnand van Stam, "Tivo: Making show recommendations using a distributed collaborative filtering architecture," Proceedings of the
-

- 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.394-401, 2004.
- [69] Alfonso Martinez, Jose J. Pazos-Arias, Ana Fernández Vilas, Jorge García-Duque and Martín López-Nores, "What's on TV Tonight? An Efficient and Effective Personalized Recommender System of TV Programs," IEEE Transactions on Consumer Electronics, Vol.55, No.1, pp.286-294, 2009.
- [70] Paul Cotter and Barry Smyth, "Personalization Technologies for the Digital TV World," Proceedings of Prestigious Applications of Intelligent Systems (PAIS'00) Berlin, Germany: IOS, pp.701-705, 2000.
- [71] Sean M. McNee, John Riedl and Joseph A. Konstan, "Accurate Is Not Always Good : How Accuracy Metrics Have Hurt Recommender Systems," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.1097-1101, 2006.
- [72] Sean M. McNee, John Riedl and Joseph A. Konstan, "Making Recommendations Better : An Analytic Model for Human-Recommender Interaction," Proceedings of ACM Special Interest Group on Computer-Human Interaction (ACM SIGCHI), pp.1103-1108, 2006.
- [73] 清水拓也, 土方嘉徳, 西田正吾, "発見性を考慮した協調フィルタリングアルゴリズム," 電子情報通信学会論文誌, Vol.J91-D, No.3, pp. 538-550, 2008.
- [74] Mukund Deshpande and George Karypis, "Item-based top-N recommendation algorithms," Journal of ACM Transactions on Information Systems, Vol.22, No.1, pp.143-177, 2004.
- [75] Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan and Georg Lausen, "Improving Recommendation Lists Through Topic Diversification," Proceedings of World Wide Web Conference, pp.22-32, 2005.
- [76] 村上知子, 森紘一郎, 折原良平, "推薦の意外性向上のための手法とその
-

- 評価," 人工知能学会論文誌, Vol.24, No.5, pp.428-436, 2009.
- [77] James Davidson, Benjamin Liebald, Junning Liu, Palash Nandy, Taylor Van Vleet, Ullas Gargi, Sujoy Gupta, Yu He, Mike Lambert, Blake Livingston and Dasarathi Sampath, "The YouTube Video Recommendation System," Proceedings of the fourth ACM conference on Recommender systems, pp. 293-296, 2010.
- [78] Nana Yaw Asabere, "Towards a Perspective of Hybrid Approaches and Methodologies in Recommender Systems," International Journal of Computer Science and Telecommunications," Vol.3, No.11, pp.23-32, 2012.
- [79] Takayuki Akiyama, Kiyohiro Obara and Masaaki Tanizaki, "Proposal and Evaluation of Serendipitous Recommendation Method Using General Unexpectedness," Proceedings of the Workshop on the Practical Use of Recommender Systems, Algorithms and Technologies (PRSAT 2010), pp.3-10, 2010.
- [80] 田野俊一, 木根智也, 石谷規彦, "広域的強化学書によるデジタル情報推薦の活性化手法," 電気学会論文誌, Vol.121, No.7, pp.1237-1245, 2001.
- [81] 板倉豊和, "タイムシフトマシンに適したパーソナル番組推薦技術," 東芝レビュー, Vol.68, No.7, pp.58-59, 2013.
- [82] 平田直之, 猪熊広哲, 関崇秀, 長野裕一, 石井 啓之, "動画共有サイトにおけるタグ情報を利用したレコメンデーションシステムの提案," 電子情報通信学会通信ソサイエティ大会, B-7-91, p151, 2006.
- [83] Greg Linden, Brent Smith and Jeremy York, "Amazon.com Recommendations:Item-to-Item Collaborative Filtering," IEEE Internet Computing, Vol.7, No.1, pp76-80, 2003.
- [84] 金田瑞規, 渡辺裕, "モバイル環境におけるユーザクラスタリングを用いた情報推薦システムの検討," 情報処理学会研究報告, 2004(25), pp.19-24, 2004.
-

- [85] Adomavicius, G. and Youngok Kwon, "New Recommendation Techniques for Multicriteria Rating Systems," IEEE Intelligent Systems, Vol.22, No.3, pp.48-55, 2007.
- [86] Twitter, Inc., Twitter : <https://twitter.com/>
- [87] Facebook, Inc., Facebook : <https://www.facebook.com/>
- [88] ANNUAL REPORT 2013 :
https://materials.proxyvote.com/Approved/90184L/20140328/AR_202076/pubData/source/Twitter,%20Inc%202013%20Annual%20Report.pdf
- [89] Facebook Reports First Quarter 2014 :
<http://investor.fb.com/releasedetail.cfm?ReleaseID=842071>
- [90] LINE公式ブログ : <http://official-blog.line.me/ja/archives/1001168643.html>
- [91] 総務省, 情報通信白書, 2013.
- [92] (株)インプレスビジネスメディア, ソーシャルコマース調査報告書2013, 2012.
- [93] Dave Kerpen, *Likeable Social Media*, McGraw-Hill, 2011.
- [94] Stephen D. Rappaport, *Listen First!*, 翔泳社, 2012.
- [95] 梅島彩奈, 宮部真衣, 荒牧英治, 灘本明代, "災害時Twitterにおけるデマとデマ訂正RTの傾向," 情報処理学会研究報告, データベース・システム研究会報告 2011-DBS-152(4), pp.1-6, 2011.
- [96] 保科一明, 武田利浩, 平仲幸雄, "口コミ型情報推薦システムの提案と評価," 情報処理学会研究報告, 10-6-A3-3, pp.1-4, 2010.
- [97] Mohsen Jamali and Martin Ester, *Mining social networks for recommendation*, Tutorial at the 11th IEEE International Conference on Data Mining (ICDM) Vancouver, Canada, 2011.
- [98] Ching-man Au Yeung and Tomoharu Iwata, "Strength of social influence in trust networks in product review sites," Proceedings of the fourth ACM international conference on Web search and data mining, pp.495-504, 2011.

- [99] 福島良典, 大澤幸生, "ソーシャルメディアを利用したセレンディピティな情報推薦," The 26th Annual Conference of the Japanese Society for Artificial Intelligence 2012, 3E1-R-6-6, pp.1-4, 2012.
- [100] 澤井里枝, 有安香子, 藤沢寛, 金次保明, "SNSを利用した協調フィルタリングによる番組推薦手法," 情報処理学会研究報告, 2010-DBS-151(43), pp.1-8, 2010.
- [101] 志甫谷匠, 中島伸介, 角谷和俊, "トレンド分析および対象グループ粒度推定に基づく情報推薦システムの提案," DEIM Forum 2010, A3-5, 2010.
- [102] 安原智樹, マーケティングの基本, 日本実業出版社, 2009.

著者略歴

廣 井 和 重（ひろい かずしげ）

1988 年 4 月 上智大学 理工学部 機械工学科 入学
1992 年 3 月 上智大学 理工学部 機械工学科 卒業
1992 年 4 月 上智大学大学院 理工学研究科 博士前期課程 入学
1994 年 3 月 上智大学大学院 理工学研究科 博士前期課程 修了
1994 年 4 月 株式会社 日立製作所 入社
2013 年 4 月 電気通信大学大学院 情報システム学研究科 博士後期課程 入学
2015 年 6 月 電気通信大学大学院 情報システム学研究科 博士後期課程 修了